

From: Ibrahim, Medina A.
Sent: Monday, September 12, 2005 11:02 AM
To: STIC-Biotech/ChemLib
Subject: 10/713,381

Please search the following:

1. SEQ ID NO: 1 and 2.
2. the sequence from nucleotide position 1 to nucleotide position 1311 of SEQ ID NO: 1 or 2.
3. the sequence from nucleotide position 1155 to nucleotide position 1311 of SEQ ID NO: 1 or 2.
4. the sequence from nucleotide position 1179 to nucleotide position 1208 of SEQ ID NO: 1 or 2.
5. the sequence from nucleotide position 1239 to nucleotide position 1278 of SEQ ID NO: 1 or 2.
6. an oligo of SEQ ID NO: 1 or 2.

Please search both commercial and patent (including pending) databases. Thanks

Medina A. Ibrahim
Patent Examiner, GAU-1638
transgenic plants and plant breeding
Remsen-2B07
mailbox-2C18
(571) 272-0797

RECEIVED
SEP 12 2005
STIC-BIOTECH/CHEM. DIVISION
(STIC)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 9/13/05
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 17 AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable:

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This page blank (uspto)

Qy	1	CCATGGTGTCTCTATGAAAAAGATAGTACCAATGTGTCTATATCCGTTTCTTAGGGTCC	60
Db	1	CCATGGTGTCTCTATGAAAAAGATAGTACCAATGTGTCTATATCCGTTTCTTAGGGTCC	60
Qy	61	CTTCTTCTGCTTATTACTGACTGAATGGGGTTTCAAAAACTTCCACGGGTGCATGAT	120
Db	61	CTTCTTCTGCTTATTACTGACTGAATGGGGTTTCAAAAACTTCCACGGGTGCATGAT	120
Qy	121	CTGCATGTTTCCACTTCTCCCACTGCGGTGACATTTCTTGATGTGCGTGTCCCAT	180
Db	121	CTGCATGTTTCCACTTCTCCCACTGCGGTGACATTTCTTGATGTGCGGTGTCCCAT	180
Qy	181	CTACCGGAGCCCATCAGACACTTTGGGACACCCATCAAGGGCTTTGGATGGGCCA	240
Db	181	CTACCGGAGCCCATCAGACACTTTGGGACACCCATCAAGGGCTTTGGATGGGCCA	240

Db	181	CTGACCGAGGCCCATGAGACACCTTTCCGGGACACCACCATCAAGGGCCTTTCCGATGGCCCA	240
Qy	241	CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCACAATCGTCACTATA	300
Db	241	CGAGACGTATCGGGTCGTGGTGATCCAGGGGATATATGTCCCCACAATCGTCACTATA	300
Qy	301	TTATATTTCTTTAGATTTTATTTAATTTTTTGGAAAAATAACCAATTTATCTTTGGTGA	360
Db	301	TTATATTTCTTTAGATTTTATTTAATTTTTTGGAAAAATAACCAATTTATCTTTGGTGA	360
Qy	361	GGGACCTCAGCATAGATTTTTCGCTTAGGGCCACAAATGCGAGGACGAGCATCTAGTG	420
Db	361	GGGACCTCAGCATAGATTTTTCGCTTAGGGCCACAAATGCGAGGACGAGCATCTAGTG	420
Qy	421	TCACATAATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATACACT	480
Db	421	TCACATAATTTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATACACT	480
Qy	481	CGAAAGCTATCATGTATGTATTTAAAGAAACATTTAATAACGACATCTCTTTAAAAA	540
Db	481	CGAAAGCTATCATGTATGTATTTAAAGAAACATTTAATAACGACATCTCTTTAAAAA	540
Qy	541	CAAGCATATTTGGAAGAGACAAATATGTTACAGTTTCAAAACATCTAAGACGACAA	600
Db	541	CAAGCATATTTGGAAGAGACAAATATGTTACAGTTTCAAAACATCTAAGACGACAA	600
Qy	601	TTATATCGAAAGGTAGACTATGACGTTCCAGATTTTCTTTTCATTTCTTTATTTGTT	660
Db	601	TTATATCGAAAGGTAGACTATGACGTTCCAGATTTTCTTTTCATTTCTTTATTTGTT	660
Qy	661	ATTGTTTATATATACATTTTCTTCTCTTACATAGGTATTTCTTCCGATTTTATAA	720
Db	661	ATTGTTTATATATACATTTTCTTCTCTTACATAGGTATTTCTTCCGATTTTATAA	720
Qy	721	ATGACATAAAGCATTTTATTAATAAGACGACATGCGTGAATCTCGTCCAAAAATC	780
Db	721	ATGACATAAAGCATTTTATTAATAAGACGACATGCGTGAATCTCGTCCAAAAATC	780
Qy	781	TTTCTGATTTTTTTTAAAGCTAGTTTGGCAACCCCTGTTCTTTCAAAGAAATTTGATTT	840
Db	781	TTTCTGATTTTTTTTAAAGCTAGTTTGGCAACCCCTGTTCTTTCAAAGAAATTTGATTT	840
Qy	841	TTTCAAAAAAATTAAGTTTATTTTCTTTTATATAATAGAAAACCTTAGAAAAATAGAT	900
Db	841	TTTCAAAAAAATTAAGTTTATTTTCTTTTATATAATAGAAAACCTTAGAAAAATAGAT	900
Qy	901	TGCCAGACTAGGCCCTAGAAATGTTTTCCAAATAATTAACAATCAGCTGTATATAATTTG	960
Db	901	TGCCAGACTAGGCCCTAGAAATGTTTTCCAAATAATTAACAATCAGCTGTATATAATTTG	960
Qy	961	GCCAGCCCCATAAATTAATTTAAACCAGAACTGAAATCGAGCGAAACCAATCTGACAT	1020
Db	961	GCCAGCCCCATAAATTAATTTAAACCAGAACTGAAATCGAGCGAAACCAATCTGACAT	1020
Qy	1021	TTCTCTAGTATGTAAAAAGGAGAGAGAGAGAGAAATCAAGTTTAAGTCAATGTCCC	1080
Db	1021	TTCTCTAGTATGTAAAAAGGAGAGAGAGAGAGAAATCAAGTTTAAGTCAATGTCCC	1080
Qy	1081	TGAGATGTGCGGGTTTGGCAACGATGCGCACCGTAATCATAGCTCATAGGTGCGCTTAGTGA	1140
Db	1081	TGAGATGTGCGGGTTTGGCAACGATGCGCACCGTAATCATAGCTCATAGGTGCGCTTAGTGA	1140
Qy	1141	GGTTTGGCGAGCTCTCGTGCATCTCCATGCGCATCTACATGCTGTTTCAACCGTTTCGTC	1200
Db	1141	GGTTTGGCGAGCTCTCGTGCATCTCCATGCGCATCTACATGCTGTTTCAACCGTTTCGTC	1200
Qy	1201	TTGTTTCATGTTCCAAAGCCTTGCTATTTCTGAACCAAGAGATACCTACTCCCAACAAT	1260
Db	1201	TTGTTTCATGTTCCAAAGCCTTGCTATTTCTGAACCAAGAGATACCTACTCCCAACAAT	1260
Qy	1261	CCATCTTACTCATGCAACTTCCATGCAACGACCATATGTGTTCTGTGAC 1311	
Db	1261	CCATCTTACTCATGCAACTTCCATGCAACGACCATATGTGTTCTGTGAC 1311	

AX224395	LOCUS	AX224395	1394 bp	DNA	linear	PAT 10-SEP-2001
AX224395	DEFINITION	Sequence 2 from Patent WO0160997.				
AX224395	ACCESSION	AX224395				
AX224395.1	VERSION	GI:15554637				
	KEYWORDS					
SOURCE	ORGANISM	Zea mays				
		Zea mays				
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	AUTHORS	1 Albertsen, M. C., Fox, T. W., Garnaat, C. W., Huffman, G. and Kendall, T. L.				
JOURNAL	TITLE	Male tissue-preferred regulatory region and method of using same				
	PATENT	Patent: WO 0160997-A 2 23-AUG-2001.				
FEATURES	source	PIONEER HI-BRED INTERNATIONAL, INC. (US) Location/Qualifiers 1..1394 /organism="Zea mays" /mol_type="unassigned DNA" /db_xref="taxon:4579"				
ORIGIN						
Query Match		100.0%; Score 1311; DB 6; Length 1394;				
Best Local Similarity		100.0%; Pred. No. 3,6e-252;				
Matches 1311, Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	CCATGGTGTCCTATGAAAAAGAGATCAAAATGTCCTATATCCGTTTCTTAGGGTCC	60			
DB	1	CCATGGTGTCCTATGAAAAAGAGATCAAAATGTCCTATATCCGTTTCTTAGGGTCC	60			
QY	61	CTTCTTTCGCTTATATCTGATGATCGGGGTATCAAAAACCTTCCACGGGTGCATGAT	120			
DB	61	CTTCTTTCGCTTATATCTGATGATCGGGGTATCAAAAACCTTCCACGGGTGCATGAT	120			
QY	121	CTCATGTTCCACTTCTTCCCACTCGCGGTGACATTTCTTGATGTCGGTGTCCCAT	180			
DB	121	CTCATGTTCCACTTCTTCCCACTCGCGGTGACATTTCTTGATGTCGGTGTCCCAT	180			
QY	181	CTGACCGAGGCGCATCGACACCTTTGGGGACACCATCAAGGACCTTTGGATGGGCCA	240			
DB	181	CTGACCGAGGCGCATCGACACCTTTGGGGACACCATCAAGGACCTTTGGATGGGCCA	240			
QY	241	CGAGACGATCGGGTGTGTGATCGAGGGATATATGTCCCCCAATCGTCACTATA	300			
DB	241	CGAGACGATCGGGTGTGTGATCGAGGGATATATGTCCCCCAATCGTCACTATA	300			
QY	301	TTATTTATTTCTTATATATATTTAATTTTGGAAAAATTAACAACTTATCTTTGTGTA	360			
DB	301	TTATTTATTTCTTATATATATTTAATTTTGGAAAAATTAACAACTTATCTTTGTGTA	360			
QY	361	GGGCGTCGACATATATTTTGGCTTAAAGGACCGAGAAATGCGAGACAGCATGTCTAGTG	420			
DB	361	GGGCGTCGACATATATTTTGGCTTAAAGGACCGAGAAATGCGAGACAGCATGTCTAGTG	420			
QY	421	TCACATATTTGGCATCTACCCAGAACAAATTTAAAAAATTAACAACTTATCTTCACT	480			
DB	421	TCACATATTTGGCATCTACCCAGAACAAATTTAAAAAATTAACAACTTATCTTCACT	480			
QY	481	CGAAGGATCATATATATATTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAA	540			
DB	481	CGAAGGATCATATATATATTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAAA	540			
QY	541	CAAGCATATTTTGGAAAGAGCAAAATTTATGTTACAGTTTCAAAATCTTAAGAGGACAA	600			
DB	541	CAAGCATATTTTGGAAAGAGCAAAATTTATGTTACAGTTTCAAAATCTTAAGAGGACAA	600			
QY	601	TTATATGAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660			
DB	601	TTATATGAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660			
QY	601	TTATATGAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660			
DB	601	TTATATGAAAGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	660			

QY 661 ATTGTTTATATACATTTCTCTTACATAGATGATTTTCTCCGATTTTATAA 720
 DB 661 ATGTTTATATACATTTCTCTTACATAGATGATTTTCTCCGATTTTATAA 720
 QY 721 ATGACTATTAAGTCATTTTATATAGAGCAGCATGCTGATGATCTCGTCAAAATC 780
 DB 721 ATGACTATTAAGTCATTTTATATAGAGCAGCATGCTGATGATCTCGTCAAAATC 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAACCTTAGAAAAATAGACT 900
 DB 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAACCTTAGAAAAATAGACT 900
 QY 901 TGGCAGACTAGCCCTAGATATGTTTCCCAATTAATCACTGTGTATTAATTTG 960
 DB 901 TGGCAGACTAGCCCTAGATATGTTTCCCAATTAATCACTGTGTATTAATTTG 960
 QY 961 GCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 DB 1021 TTCTCTAGATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 QY 1081 TGAGATGTGGGTTTGGCAAGATAGCCACCGTATCATAGCTATGATGCTTCAAGTCA 1140
 DB 1081 TGAGATGTGGGTTTGGCAAGATAGCCACCGTATCATAGCTATGATGCTTCAAGTCA 1140
 QY 1141 GGTTCGGACGCTCTGTCATCTCACATGSCATCTACATGCTTGTCAACCGTTCGTC 1200
 DB 1141 GGTTCGGACGCTCTGTCATCTCACATGSCATCTACATGCTTGTGTCAACCGTTCGTC 1200
 QY 1201 TTGTTCCATGCTGCAAGCCCTTGCCTATCTGAAACCAAGAGATACCTACCTCCAAACAT 1260
 DB 1201 TTGTTCCATGCTGCAAGCCCTTGCCTATCTGAAACCAAGAGATACCTACCTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACGACCATATGTTTCTCGAAC 1311
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACGACCATATGTTTCTCGAAC 1311
 RESULT 3
 LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
 DEFINITION Male tissue-preferred regulatory region and method of using same.
 ACCESSION BD062176 GI:22607781
 VERSION 1
 KEYWORDS JP 2001520523-A/1.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1394)
 AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patient: JP 2001520523-A 1 30-OCT-2001;
 COMMENT PIONEER HI BRED INTERNATIONAL INC
 PN JP 2001520523-A/1
 PD 30-OCT-2001
 PF 19-JUN-1998 JP 1999504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN,
 PI TIMMY L KENDALL
 PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
 PC A01H5/00
 CC Strandedness: Single;

CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 Source Location/Qualifiers
 1.1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1311; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 3.6e-252;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAAAGGCTC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAAAGGCTC 60
 QY 61 CTTCCTGCTGCTATTAAGTGAATCGGGGTTAAGAAAACTTCACGGGGTGCATGAT 120
 DB 61 CTTCCTGCTGCTATTAAGTGAATCGGGGTTAAGAAAACTTCACGGGGTGCATGAT 120
 QY 121 CTCATGTTTCACTTCTCCACCTCGGTTGACATTTCTGGATGCTGGTCCAT 180
 DB 121 CTCATGTTTCACTTCTCCACCTCGGTTGACATTTCTGGATGCTGGTCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACTTTTGGGACACCCATCAGAGGCTTTGGATGCCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACTTTTGGGACACCCATCAGAGGCTTTGGATGCCCCA 240
 QY 241 CGAGAGTATCGGGGTGCTGTATCGAGGGGATATATGTCCTCCCAATGCTACCTAT 300
 DB 241 CGAGAGTATCGGGGTGCTGTATCGAGGGGATATATGTCCTCCCAATGCTACCTAT 300
 QY 301 TTATATCTCTTATGATATTTATTTTGGAAAAATACAACTTATACCTTTGTGTA 360
 DB 301 TTATATCTCTTATGATATTTATTTTGGAAAAATACAACTTATACCTTTGTGTA 360
 QY 361 GGGCTCAGATGATGATTTTTCCTTAAAGGCTTCAAGAGACAGCATGCTATGATG 420
 DB 361 GGGCTCAGATGATGATTTTTCCTTAAAGGCTTCAAGAGACAGCATGCTATGATG 420
 QY 421 TCCACTATTTGGGACCTCCAGAACAGATTTTAAAAAATTAACCAAGTAACTTACCT 480
 DB 421 TCCACTATTTGGGACCTCCAGAACAGATTTTAAAAAATTAACCAAGTAACTTACCT 480
 QY 481 CGAAAGCTATCATGATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAA 540
 DB 481 CGAAAGCTATCATGATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAA 540
 QY 541 CAAGCATATTTGAAAGAGACAAATATGTTTACAGTTTACAAACATCTAAGAGCGACA 600
 DB 541 CAAGCATATTTGAAAGAGACAAATATGTTTACAGTTTACAAACATCTAAGAGCGACA 600
 QY 601 TTATATCGAAGGTATGATGATGCTGATGATTTTCTTTTCACTCTGATTTTGT 660
 DB 601 TTATATCGAAGGTATGATGATGCTGATGATTTTCTTTTCACTCTGATTTTGT 660
 QY 661 ATTGTTTATATACATTTCTCTTACATAGAGATTTTCTCCGATTTTATAA 720
 DB 661 ATTGTTTATATACATTTCTCTTACATAGAGATTTTCTCCGATTTTATAA 720
 QY 721 ATGACTATTAAGTCATTTTATATAGAGCAGCATGCTGATGATCTCGTCAAAATC 780
 DB 721 ATGACTATTAAGTCATTTTATATAGAGCAGCATGCTGATGATCTCGTCAAAATC 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAACCTTAGAAAAATAGACT 900
 DB 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAACCTTAGAAAAATAGACT 900

QY	901	GGCCGAGCTAGCCCTNAGATGTTTTCCCAATTAATTAATCACTGTGTAAATTAATTTTG	960		
Db	901	TGCCAGACTTAGCCCTNAGATGTTTTCCCAATTAATTAATCAATCACTGTGTAAATTAATTTTG	960		
QY	961	GCCAGCCCCCATTAATTAATTTAAACCGAAATCGAAATCGAGGAAACCAATCTGAGCTAT	1020		
Db	961	GCCAGCCCCCATTAATTAATTTAAACCGAAACTGMAATCGAGGAAACCAATCTGAGCTAT	1020		
QY	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCATTTGCC	1080		
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCATTTGCC	1080		
QY	1081	TGAGATGTGCGGTTTGCGACAGATAGCCACCGTATCTACTGATAGTGTCCTACGCTCA	1140		
Db	1081	TGAGATGTGCGGTTTGCGACAGATAGCCACCGTATCTACTGATAGTGTCCTACGCTCA	1140		
QY	1141	GGTTGGGAGGCTCTGTGTCATCTCAATGGATATCTACGCTTTGTTCAACGGTTGCTC	1200		
Db	1141	GGTTGGGAGGCTCTGTGTCATCTCAATGGATATCTACGCTTTGTTCAACGGTTGCTC	1200		
QY	1201	TTGTTCCATCGTCCAGGCTTGCCCTATTCTGAACCAAGAGATACCTTCCCAACAT	1260		
Db	1201	TTGTTCCATCGTCCAGGCTTGCCCTATTCTGAACCAAGAGATACCTTCCCAACAT	1260		
QY	1261	CCATCTTACTCATGCAACTTCCATGCAACAACGACATATGTTTCTGTAAC	1311		
Db	1261	CCATCTTACTCATGCAACTTCCATGCAACAACGACATATGTTTCTGTAAC	1311		
RESULT 4					
BD062177		1394 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD062177				
DEFINITION		Male tissue-preferred regulatory region and method of using same.			
ACCESSION	BD062177				
VERSION	BD062177.1	GI:22607782			
KEYWORDS		JP 2001520523-A/2.			
SOURCE		Homo sapiens (human)			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS		1 (baaes I to 1394)			
		Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.			
TITLE		Male tissue-preferred regulatory region and method of using same			
JOURNAL		Patent: JP 2001520523-A 2 30-OCT-2001;			
		PIONEER HI BRED INTERNATIONAL INC			
COMMENT		PN JP 2001520523-A/2			
		PD 30-OCT-2001			
		PF 19-JUN-1998 JP 1999504910			
		PR 23-JUN-1997 US 08/880499			
		PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,			
		PI TIMMY L KENDALL			
		PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC			
		PC C07K14/34,C12Q1/68,			
		PC A01H5/00			
		CC Strandedness: Single;			
		CC Topology: Linear;			
		FH Key			
FEATURES					
source		Location/Qualifiers.			
		1..1394			
		/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
ORIGIN					

Query Match	100.0%	Score 1311;	DB 6;	Length 1394;
Best Local Similarity	100.0%	Prod. No. 3.6e-25;		
Matches 1311; Conservative	0;	Mismatches	0;	Gaps 0;

Qy 1 CCATGGTGTCTCATGAAAAAGAGATCATATGTCATATCCGTTTTCTTAGGGTCC 60

Db 1 CCATGGTGTCTCATGAAAAAGAGATCATATGTCATATCCGTTTTCTTAGGGTCC 60

QY	61	CTTCTCTTGCCTTATTACTGACTGAATGGGGGTTACAAAAAACTTCCACGGGTGCATGAT	120
Db	61	CTTCTCTTGCCTTATTACTGACTGAATGGGGGTTACAAAAAACTTCCACGGGGTGCATGAT	120
QY	121	CTCCATGTTCCACTCTTCCCACTCGGGCGTTGACATTTCTTGGATGTCGGTGGTTCCTCAT	180
Db	121	CTCCATGTTCCACTCTTCCCACTCGGGCGTTGACATTTCTTGGATGTCGGTGGTTCCTCAT	180
QY	181	CTGACCGAGGCCCATCAGACACTTTTGGGACACCATTCAGAGGCGCTTTCCGATGGCCCA	240
Db	181	CTGACCGAGGCCCATCAGACACTTTTGGGACACCATTCAGAGGCGCTTTCCGATGGCCCA	240
QY	241	CGAGACGTAATGGGGTCGGTGGTGAATCCAGAGGATATATGTCCCCCAATTCGTCACTATA	300
Db	241	CGAGACGTAATGGGGTCGGTGGTGAATCCAGAGGATATATGTCCCCCAATTCGTCACTATA	300
QY	301	TTATATATCTTATAGATATTTATTTATTTTGGAAAAATTAACAAACTTATACTTTTGTGTGA	360
Db	301	TTATATATCTTATAGATATTTATTTATTTTGGAAAAATTAACAAACTTATACTTTTGTGTGA	360
QY	361	GGGCGCTCAGATAGATTTTTCGTTTAGGGCCCGAAGATGGAGAGCAAGCATGTCTAGTG	420
Db	361	GGGCGCTCAGATAGATTTTTCGTTTAGGGCCCGAAGATGGAGAGCAAGCATGTCTAGTG	420
QY	421	TTCAATATTGGGACATACCCAGAAACAGATTTAAAAAAATTAACCAAGTAACTAATCCACT	480
Db	421	TTCAATATTGGGACATACCCAGAAACAGATTTAAAAAAATTAACCAAGTAACTAATCCACT	480

Qy	481	CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAACGATCCTTCTTAAAAA	540
Dh	481	CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAACGATCCTTCTTAAAAA	540

[illegible]

601 TTATATCGAAGGTAGCTATGACGTTCAATTTTCTTTTCATCTTGTTATTTTGGTT 660

Db 601 TTAATCGAAGGTAGCTATGACGTTGAGATTTTCTTTTCATCTCTGTATTTGTG 660

Qy 661 ATGTGTTTAAATACATTTTCTTCTTACAAATAGAGTGAATTTCTCCGATTTTATAAA 720

Db 661 ATTGTTTATATACATTTTCTTCTTACAATGAGTGATTTTCTTCCGATTTTATPAA 720

Ov 721 ATGACTATAAGTCATTTTATATPAGAGCAGCGATGTCGTAGATTTCTGTTCAAAAATC 780

Db 721 ATGACTATAAGTCATTTTATATATAGAGCAGCATGTGCTAGATTCTCGTTCAAAATC 780

Db 781 TTTTCGATTTTTTTAAGAGCTAGTTGGCAACCCGTTTCTTCAAGAATTTTGATTTT 840

Qy	841	TTCAAAAAAATGATTATTTTCTCTTATAAATAGAAAACACTAGAAAAATGAGT	900
Db	841	TTCAAAAAAATGATTATTTTCTCTTATAAATAGAAAACACTAGAAAAATGAGT	900

QY 901 TGCCAGACTAGCCCTAGAA TGT TTTCCCATTAATTAATTAATTTG 960

Dh 901 TGCCAGACTAGCCCTAGAA TGT TTTCCCATTAATTAATTAATTTG 960

961 GCGAGCCCATTAATTATTTAAACCGAAGTGAATCGAGCGAACCAGTCTGAGCTAT 102

[illegible]

Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATTCAGTTTTTAAGTCATGTCC	108
Qy	1081	TGAGATGTGCGGTTTGGCAACGATATGCCACCGTAAATCATATAGTCTATAGTGCCCTACGTCA	114

Db 1081 TGAATGTCGGTTGGCAACGATGACCCTGATCATAGCTCTATAGGTCCTTAGCTCA 1144

Db	1141	GGTTGGCAGCTCTGTGTATCTACATGCGATCTACTATGCTTGTTCACACCGTTGCCTC	12000
Oy	1201	TTGTTCCATCGTCGAAGCCCTTGCCATTCTCGAACCAAGAGATACCTACTGCCAACAAT	12600
Db	1201	TTGTTCCATCGTCGAAGCCCTTGCCATTCTCGAACCAAGAGATACCTACTGCCAACAAT	12600
Oy	1261	CCATTTTCTCATGCAACTCTCCATCGCAACGCAACGCAATATGTTCTCTGAC	1311
Db	1261	CCATTTTCTCATGCAACTCTCCATCGCAACGCAACGCAATATGTTCTCTGAC	1311
RESULT 5			
AF360356			
LOCUS	AF360356	3343 bp	DNA linear PLN 12-MAY-2001
DEFINITION	Zea mays male fertility protein (Ms45) gene, complete cds.		
VERSION	AF360356		
KEYWORDS	AF360356.1 GI:14028756		
SOURCE	.		
ORGANISM	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD		
	clade; Panicoidae; Andropogoneae; Zea.		
	1 (bases 1 to 3343)		
REFERENCE	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
AUTHORS	Cloning of Ms45, a gene required for male fertility from Zea mays		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3343)		
AUTHORS	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
TITLE	Direct Submission		
REFERENCE	Submitted (13-MAR-2001) Trait and Technology Development, Pioneer		
AUTHORS	Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,		
JOURNAL	IA 50131-1004, USA		
FEATURES			
source	location/Qualifiers		
	1..3343		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/culivar="B73"		
	/db_xref="taxon:4577"		
	/chromosome="9L"		
	<1392..>2942		
	/gene="Ms45"		
	join(1392..1768,1898..2182,2280..2447,2534..>2942)		
	/product="male fertility protein"		
	join(1392..1768,1898..2182,2280..2447,2534..>2942)		
	/gene="Ms45"		
	join(1392..1768,1898..2182,2280..2447,2534..>2942)		
	/product="male fertility protein"		
	join(1392..1768,1898..2182,2280..2447,2534..>2942)		
	/gene="Ms45"		
	/codon_start=1		
	/product="male fertility protein"		
	/protein_id="AAK52489.1"		
	/db_xref="GI:14028757"		
	/translation="MEKNILQMRGRDGIVOYPHLPFAALALLVADPRLSLAIV		
	DYRVYKELAPYEVMGSWPRDNASRLRGLSEVGVEFGESIEPLQGRPGAYGLA		
	DGRVVRKMGEAEAGMETPAVNPNDSBEVCAGAVSTRKHKEKEFCGRPLGLEFHEE		
	TGELYVNADAYGLMNVVQSGGVASSVAAREADGDIRANDLDVHRNDSVFETDTSMY		
	RKHQLNTLLEEGEGRLRLRPDPERSGVHVLKGLVINGVQISEDHQFLIFSETTNC		
	RIKMWLEGPAGEVEVPANLPSPDNVRSNGROFPVALIDCCRTPAOEYAKAPMLR		
	TLVKRPPLSKVLTMKARBMHTLVALLDGSRGVVEVLEDRGHEVMKLVSIVREVRGKK		
	LWIGTVAMNHAIITPYLED"		
ORIGIN			
	Query Match	100.0%; Score 1311; DB 8; Length 3343;	
	Best Local Similarity	100.0%; Pred. No. 3.5e-252;	
	Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Oy	1	CCATGCGTCTCTATGAAAAAGATGAGTAACAATGCTATATCCGTTTTCTTAGGCTC	60
Db	1	CCATGCGTCTCTATGAAAAAGATGAGTAACAATGCTATATCCGTTTTCTTAGGCTC	60
Oy	61	CTTCTTCGCTTATTAAGAATAAGGGGGTTTCAAAAACCTTCCACGGGTGCATGAT	120
Db	61	CTTCTTCGCTTATTAAGAATAAGGGGGTTTCAAAAACCTTCCACGGGTGCATGAT	120

QY	121	CTCCATGTTCCACTTCTCCACCTCGCGTGCACATTTCTTGATGTCGGTGGTTCCCAT	180
Db	121	CTCCATGTTCCACTTCTCCACCTCGCGTGCACATTTCTTGATGTCGGTGGTTCCCAT	180
QY	181	CTGACCGGAGGCCAATCAGACACCTTTGGGGGACCCCATCAAGGGGCTTTGGGATGGGCCA	240
Db	181	CTGACCGGAGGCCAATCAGACACCTTTGGGGGACCCCATCAAGGGGCTTTGGGATGGGCCA	240
QY	241	CGAAGCGTATCGGGTCGTGTGATCCAGGGGATATGTCCCGCACATCGTCACTATA	300
Db	241	CGAAGCGTATCGGGTCGTGTGATCCAGGGGATATATGTCCCGCACATCGTCACTATA	300
QY	301	TTATTTATCTTTAGATATATTATTTAATTTTGGAAAAATACAAACTTATACCTTTGTGTGA	360
Db	301	TTATTTATCTTTAGATATATTATTTAATTTTGGAAAAATACAAACTTATACCTTTGTGTGA	360
QY	361	GGGGCTCAGCATAGATTTTGGCTTAAGGGGCCAGAAATGCGAGGACCGAGCCATGTCATGTG	420
Db	361	GGGGCTCAGCATAGATTTTGGCTTAAGGGGCCAGAAATGCGAGGACCGAGCCATGTCATGTG	420
QY	421	TCCACTATGTGCACCTACCCAGAACAGATTTTAAAAAAAATAACAAAGTAACCTATCCACT	480
Db	421	TCCACTATGTGCACCTACCCAGAACAGATTTTAAAAAAAATAACAAAGTAACCTATCCACT	480
QY	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTCTTAAAAAA	540
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTCTTAAAAAA	540
QY	541	CAACCATATTTCCGAAAGAGCAAAATTTATNGTTACAGTTTACAAACATCTAAGAGGACAAA	600
Db	541	CAACCATATTTCCGAAAGAGCAAAATTTATNGTTACAGTTTACAAACATCTAAGAGGACAAA	600
QY	601	TTATATCGAAAGTAAGCTATGACGTTCAGATTTTTCCTTTTCATTCCTGTATTTTGT	660
Db	601	TTATATCGAAAGTAAGCTATGACGTTCAGATTTTTCCTTTTCATTCCTGTATTTTGT	660
QY	661	ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGAGATTTTTCCTCCGATTTTATPAA	720
Db	661	ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGAGATTTTTCCTCCGATTTTATPAA	720
QY	721	ATGACTATAAAGTCAATTTTATATATAAGACGCAATGTCTGTAGATTTCTCGTCAAAAATC	780
Db	721	ATGACTATAAAGTCAATTTTATATATAAGACGCAATGTCTGTAGATTTCTCGTCAAAAATC	780
QY	781	TTTCTGATTTTATTAAGAGTAGTTTGGCAACCCGTGTTCTTTCAAGAAATTTTGATTTT	840
Db	781	TTTCTGATTTTATTAAGAGTAGTTTGGCAACCCGTGTTCTTTCAAGAAATTTTGATTTT	840
QY	841	TTCAAAAAAAATTAAGTTATTTTCTCTTATATAAATAGAAAACACTTAGAAAAATAGAGT	900
Db	841	TTCAAAAAAAATTAAGTTATTTTCTCTTATATAAATAGAAAACACTTAGAAAAATAGAGT	900
QY	901	TGCGACAGCTAGCCTTAGAATGTTTCCCAATAAATTAACATCACTGTGTATATATATTTG	960
Db	901	TGCGACAGCTAGCCTTAGAATGTTTCCCAATAAATTAACATCACTGTGTATATATATTTG	960
QY	961	GCGAGCCCCATTAATTTATTTAAACGAAACCTGAATGAGGGAACCAAACTGAGCTAT	1020
Db	961	GCGAGCCCCATTAATTTATTTAAACGAAACCTGAATGAGGGAACCAAACTGAGCTAT	1020
QY	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTTTAACTATTTGTGCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTTTAACTATTTGTGCC	1080
QY	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTATATCATGTATGAGTGCTTACGTCA	1140
Db	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTATATCATGTATGAGTGCTTACGTCA	1140
QY	1141	GGTTCCGAGAGCTCGTGTCATCTCAATGAGCATATCAATGCTGTGTTCAACGGTTGTC	1200
Db	1141	GGTTCCGAGAGCTCGTGTCATCTCAATGAGCATATCAATGCTGTGTTCAACGGTTGTC	1200

QY	1201	TTGTTCCATCGCCGAAGCCTTGCCATTTCTGGAACGAAGATACCTACTCCCAACAAAT	1260
Db	1201	TTGTTCCATCGCCGAAGCCTTGCCATTTCTGGAACGAAGATACCTACTCCCAACAAAT	1260
QY	1261	CCATCTTACTCATGCAACTTCCATGCAACAGCACAATATGTTTCTGTAAAC	1311
Db	1261	CCATCTTACTCATGCAACTTCCATGCAACAGCACAATATGTTTCTGTAAAC	1311
RESULT 6			
LOCUS	AX224402	255 bp	linear
DEFINITION	Sequence 9 from Patent WO0160997.		
ACCESSION	AX224402		
VERSION	AX224402.1	GI:15554644	
KEYWORDS			
SOURCE			
ORGANISM			
	Zea mays		
	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE			
AUTHORS	1	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.	
TITLE		Male tissue-preferred regulatory region and method of using same	
JOURNAL		Patent: WO 0160997-A 9 23-AUG-2001;	
FEATURES			
	PIONEER HI-BRED INTERNATIONAL, INC. (US)		
	Location/Qualifiers		
	1..255		
	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4577"		
ORIGIN			
	Query Match	12.3%;	Score 160.6; DB 6; Length 255;
	Best Local Similarity	97.6%;	Pred. No. 5.9e-22;
	Matches	163; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	1145	CGCAGCTCTCGTGTATCTCATGATGGAATACATGCTGTTCAACCGTTCGTGTG	1204
Db	5	CGCGGAGATCCGGTGTGCATCTCATGAGCATACTACATGCTGTTCAACGTTTCGTGT	64
QY	1205	TCGACGTGCCAAGCCTTGCCATTTCTGGAACGAAGATACCTACTCCCAACAAATCAT	1264
Db	65	TCGACGTGCCAAGCCTTGCCATTTCTGGAACGAAGATACCTACTCCCAACAAATCAT	124
QY	1265	CTTACTCATGCAACTTCCATGCAACAGCACAATATGTTTCTGTAAAC	1311
Db	125	CTTACTCATGCAACTTCCATGCAACAGCACAATATGTTTCTGTAAAC	171
RESULT 7			
LOCUS	AX224396	158 bp	DNA
DEFINITION	Sequence 3 from Patent WO0160997.		
ACCESSION	AX224396		
VERSION	AX224396.1	GI:15554638	
KEYWORDS			
SOURCE			
ORGANISM			
	Zea mays		
	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
REFERENCE			
AUTHORS	1	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.	
TITLE		Male tissue-preferred regulatory region and method of using same	
JOURNAL		Patent: WO 0160997-A 3 23-AUG-2001;	
FEATURES			
	PIONEER HI-BRED INTERNATIONAL, INC. (US)		
	Location/Qualifiers		
	1..158		
	/organism="Zea mays"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4577"		
ORIGIN			

Query Match	11.1%	Pred. 146:	DB 6:	Length 158:
Best Local Similarity	99.4%	Pred. No. 5e-19;		
Matches 157;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
QY	1155	CGTGTCAATTCACATGAGCATACTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC	1213	
DB	1	CGTGTCAATTCACATGAGCATACTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC	60	
QY	1214	CAAGCTTGGCTTATTTCTGAAACCAAGAGATACCTACTCTCCCAACATTCATCTTACTCAT	1273	
DB	61	CAACCTTGGCTTATTTCTGAAACCAAGAGATACCTACTCTCCCAACATTCATCTTACTCAT	120	
QY	1274	GCAACTTCGATGCAAGACGACGACATATGTTTCCGAGC	1311	
DB	121	GCAACTTCGATGCAAGACGACGACATATGTTTCCGAGC	158	
RESULT 8				
AC147602		186199 bp	DNA	linear
LOCUS	Zea mays clone ZMMBc0334A01, *** SEQUENCING IN PROGRESS ***			6
DEFINITION	ordered clones.			
AC147602				
AC147602.5	GI:51315585			
KEYWORDS	HTG; HTGS_PHASE2; HTGS_FULFILL; HTGS_ACTIVEFIN.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.			
1	(bases 1 to 186199)			
REFERENCE	Bittern, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K. and Messing, J.			
AUTHORS	Zea mays, clone ZMMBc0334A01			
TITLE	Unpublished			
JOURNAL	2 (bases 1 to 186199)			
REFERENCE	Bittern, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Ramchaka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testa, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 186199)			
AUTHORS	Bittern, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K., Messing, J., Abouelleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,			

Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachpaka, A., Ramasamy, U., Raymond, C., Reta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Tamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 17, 2004 this sequence version replaced gi:49658659.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@road.mit.edu

Bhatl, AK and Messing, J: The Plant Genome Initiative at Rutgers, Maksum Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
(<http://pgr.rutgers.edu>)
Butler, E and Ming, R: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (<http://www.genome.arizona.edu>)

Project Information
Center project name: L30003
Center clone name: 334_A.1

Consensus Information
This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ne are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository
(<http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?>). An exact list of reads used in this assembly are available at
<http://www.broad.mit.edu/annotation/plants/maize/randomclones.html>.

NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 100617: contig of 100617 bp in length
* 100618 100717: gap of unknown length
* 100718 104730: contig of 4013 bp in length
* 104731 104830: gap of unknown length
* 104831 115104: contig of 10274 bp in length
* 115105 115204: gap of unknown length
* 115205 115305: contig of 41192 bp in length
* 115306 156396: gap of unknown length
* 156397 179936: contig of 23440 bp in length
* 179937 180036: gap of unknown length
* 180037 186199: contig of 6163 bp in length.

Location/Qualifiers
1.186199
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="ZMWBBC0334A01"

FEATURES
source

ORIGIN
/clone_1lb="CHORI-201 Maize B73 BAC Mbol Library"

Query Match 5.9%; Score 77; DB 2; Length 186199;
Best Local Similarity 71.4%; Pred. No. 2.7e-05;
Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

Qy 780 CTTTCGATTTTAAAGAGCTAGTTGGCAACCCGTTCTTCAAAAGATTGATTT 839
|||||
Db 162058 CTTTCTTTACTTTTAAAGGCTAGTTGGCAACCCGTTT-TTCTAAGAAATTTTCATTT 162116
|||||

Qy 840 TTTCAAAAAATTTAGTTATTTCTTTATTAATAGAAACACCTTGAATAATAGAG 899
|||||
Db 162117 TTTCAGTGAATTTAGTTATTTCTTTAACAATAATAGAAATTTTGTATTAATAGAG 162176
|||||

Qy 900 TTGCAGACTAGCCCTAGATGTTTCCCAATTAATAGCA 940
|||||
Db 162177 TTTCCAACTTAACCTCAAGTATATATGATGACAGACAA 162217
|||||

RESULT 9
AC117267 25769 bp DNA linear INV 21-FEB-2004
LOCUS AC117267
DEFINITION Dictyostelium discoideum chromosome 2 map 5836255-5862024 strain
AX4, complete sequence.
ACCESSION AC117267 AC115597
VERSION AC117267.2 GI:42733680
KEYWORDS HTG.
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
Eukaryota; Eukaryota; Dictyostelida; Dictyostelium.

REFERENCE
1 (bases 1 to 25769)
Gloeckner, G., Eichinger, L., Szafarski, K., Pachbat, J., Dear, P., Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and Noegel, A.A.
Sequence and analysis of chromosome 2 of Dictyostelium discoideum
Nature 418 (6883), 79-85 (2002)

JOURNAL MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
2 (bases 1 to 25769)
REFERENCE
1 (bases 1 to 25769)
Baumgart, C.
JOURNAL Direct Submission
Submitted (09-APR-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 25769)
Baumgart, C.
JOURNAL Direct Submission
Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from Geneid may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(<http://genome.imb-jena.de/dictyostelium/>)
and the University Cologne, Institute for Biochemistry I
(<http://www.uni-koeln.de/dictyostelium/project.shtml>)
Funding
Agency: Deutsche Forschungsgemeinschaft (DFG).

FEATURES
source
1.25769
Location/Qualifiers
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
complement(join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="Geneid exon scores (in order of location ranges):
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1

CDS
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="ZMWBBC0334A01"

/product="hypothetical protein"
/protein_id="AA338625.1"
/db_xref="GI:42733681"
/translation="MRLOTLPLSLILRLQKSTSSNDISTLITSTNFKNDYNO
ILSKSEWFSLEYSNNKVIYEVDPFNHYLENDRLLKEFENYVAVDSFSCFL
NSKFNLEYKKEFTSIRNNKVPENIMNIESTHENCYNNVLYNNKNOEYLV
FSKQVLSFLQNGYNIKSSKNTFKIENNNDKLFSSLPFYLNKIMENOFEDY
SPLSGMDFDKIKQPKLNDIETLIDVEVNLCSNNNNNNNNNNVDETLTLGSGSG
SGSGSGSGSGSLGSGTYNHQ"
join(2957..3173,3272..5583)
/note="Geneid exon scores (in order of location ranges) :
24.88, 254.96 - GSCU_ID dd_00753"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA338626.1"
/db_xref="GI:42733682"
/translation="MEPTKYVITYTSLNLDKPKEMWTFQSEMSLTITDTIQKSFVNSG
SSSSKGFGEAVLILDTAGQESYAMQVYRTDCENIYFSDIRSSFEVSQK
OHIERVYDRDVPYIIIVGNKVDLESRRVSRADQLARSLRVPYIETSAKTSNIE
AFETLVHPTNRYVYKVVVGGGKSAIIIOFNHVEEYPTIEDSYKQVYIS
GLPPIGSLNKKSSSSSSSSSGKGLFNKIFSGKDRSPQOASSTIDRGOI
STNRLEANLYSMSNLKVEPYITGDCYCCQCNVILSRPSNLYTGDPSTWKEP
CKTSNSNILEGEIPNCDSEVYLSPESTSTTDSKBESEIYICIDVSGWGIT
EVPYLSQSEWNAKGVKASGSPYISRLCQVQS.IPTMIDRLSIQPKRVVLYTFS
DEVMYITQSNVDGPIVIAGDKLEDFDQIIEIGRSMYDKLPTASGSDPLAKIKSL
EPVOSTALGPALIVSAIASOKMLSEVICTDGVNPGALIEDLPGAOFPEKYV
KLAQNNKTINIIIGISGSHIDIGVIGKEQNTGNITIIHPLEAREIKRLQNPMTA
TJYEMETICAPTELEIKYDSKQSLSRVYQFPYNNLTJDLTLLSRNPTFEVQYR
FQIQITRTKIDVRCIRVYSAQLOATPDRFTSTSNISILAWAFIQOAKIAQOQY
MESRLHLKATKILRSLCNTDEQEEFYFVEVIREMEAPLITCINKQOQVEKATD
DEIQVYKMKVHKSFEVGGKKDIRRKEGAEINKYNYIKFT"
complement(5711..6796)
/note="Geneid exon scores (in order of location ranges) :
68.61 - GSCU_ID dd_00752"
/codon_start=1
/product="similar to Plasmodium falciparum (isolate 3D7).
Hypothetical protein"
/protein_id="AA338627.1"
/db_xref="GI:42733683"
/translation="MSRYSKSLIERYNGINTREKSPRLINNNNNNNNN
GTTTITSGANNNSTSPSTNPEVSDRVVDHSSNNQIQIVDSFNINNNNNNN
NN
KFNNNNNYFDFSTNEMITTLNSCTLEQIGTGNDYSTNTTSPSSSSSTHLCN
KSNHNNMPSLKINQIINDNNNINSINIEFQIQQSSSSSCSTYQCPIDEHIOE
QSLIOEKQKOIDDDDDDDDDSSSPPTTTTKTFCLSRPRIKKEENID
QMSBPPLKDRPSRDLIPTIANI"
join(8924..9076,9236..10573)
/note="Geneid exon scores (in order of location ranges) :
10.37, 128.60 - GSCU_ID dd_00729"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA338628.1"
/db_xref="GI:42733684"
/translation="MSSSTITPIATTTNNNSSTNGCHNSKTRTLFKNIKNIENS
PNSAPIKNSVGIIGNGPVNNNNKDKGNSDGVNFGKNLLPTSKNSMSTTKHS
SSSSSILAKNEYIEBIIKKEKMDISTGDISPLSPLRESSPISTIDESIATATK
LTPAFVAASSSSSVKTIIDVLSFKSANTKRPIDIDIEVNHOKELIVLSPYNN
NNNNNNNNNNNSOHTANGOKYPIFSPQISPKIAYAOSSSTKTNINNNNNNTPT
TKKANSNITTPQSNRNNHNSYTSKSKONPIPSLTLATPNOKNTTTTSSKSTT
TTPSKKIKINDVAAFAAVANSTTTTQPAIIDDATPTNESPKTPTLKETDPRKLA
ARQROIDIGKMTAGYKNYIALVPKSRKRPDTPPKNQOVCSKRWDOIKKWRQLH
ENYDWTGIFPNEKEVEASNNSTESTLEKMKLEISENEHNN"
complement(10972..11979)
/note="Geneid exon scores (in order of location ranges) :
23.29 - GSCU_ID dd_00728"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA338629.1"
/db_xref="GI:42733685"
/translation="MEBKIKYFVEIIDPONOKKIOEFTSKLIGKESGPTPKPIVY
EKLWNTQSIIESILKTNGTINKSIFEEBFSYCYGIGELEBEPPIVPLKFTTGGIL
DDTIFKINSLNKLEKEKULNSLIYNNKNNKILGFEPWEIENFQPTKESFOR
IINANSIWKSSIDSRNINSNPSINRCSFNEYEKRSSDASGDFILITISMIGIMDYI

ENSIESKEPOTINYHAKSFLLINDLYSPNREINENDLNTYKILAIQANSIQSLID
KTELIIIDHYLKLSSIEITLKYQNDSTVQLKQVQNSKILSGIYFAKKSKRY
N"
join(12884..13020,13172..15320)
/note="Geneid exon scores (in order of location ranges) :
2.42, 136.56 - GSCU_ID dd_00728"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Hypothetical 97.7 kDa protein"
/protein_id="AA338630.1"
/db_xref="GI:42733686"
/translation="MDNLIYDFLIRKEYNQGSFLDCLILSLINGLYISYDGLDCNN
FLIENKTYNKAISYGLNLYOYOLLEFETLLKRYKCTCKKXIIINTHNSLKKC
GNKKRGYILIIENNNNIINKDYQFNNNNNKNNNNNNNNNNNNNNNIIIEEDLNKI
TLKBERINGNGEITTTTTTSTTSSTTTMTTTTNGVDRNLTLLPROVARNYVIK
TOILEPHLNLNHNANKVPLPIQOLVDYFKKQYLOSMTLNDHDDDDADESYSDI
DDDDSYSGGNCSSSSSDSGSSSSINNSISSSSSILSNSLSLBSMTIRVNSL
PEGLSIFEEKEYNYIQRLPSPISLSIRFSYFNQRLAKGVISDNVISITGDSFNO
SLDGMKROLKHLOFGHKFOQTIMGQLPSSLTSLIDPRYSKGVIEIGSIPDSVT
LDYKNSGNOSIESFNPJPSKITRLVDFSENNONIKANDISNNYLSIHFGEHNS
DIGIKSLNSIREIKRGAPRDIIDCSITSIDCFKRNPLSMOTQTLTSDIGS
KRNQIIPGIRFHTLKSLNRYHNNQIIPADTLPPTLESLNLGYNREITYKDNBYD
CYGISNKGFGSSNSNFCVGGTNGLMEMLNNTSLKTLTNTYNRKILEVDLPNSI
ESLNGYHFNQPIGNVLPKLLKPLINSENNONIADGCIPIGLQIYIRNSMWF
INSLNPLFTKYINIIDSHL"
join(16184..16199,17272..17613,17707..19331)
/note="Geneid exon scores (in order of location ranges) :
0.08, 32.32, 227.42 - GSCU_ID dd_00726"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Spore coat protein sp87 precursor (Plj protein)"
/protein_id="AA338631.1"
/db_xref="GI:42733687"
/translation="MRIGEDCNKIRNEBCKHSRSCYIYNTPCGGEQKMAKSGT
TCTTNSCTRNSSNNQVYEVNSKCNELNDFIKITTPTEYSGDAIKERALKGKSP
QKNSGNSPTSCCPGSEICEGLSSGTSGGTSSGSSGTSGGSSGTSAGSGLT
SSGSSGNSVSCSTHCEPGEYCSMNVAICLSTGTCGLPTGTSSTAGSGLT
LCPIGHICVIEDSNGVYCNNGGTSGGTSGHDPGCDVDCPPOFHCEDQK
TACKVSPPTGSSSTSGHPPDCQVTPDGHGCKQKTKACVPSPTGSSSTSG
NTNPSNNCNDGPFCECKDKTKACVSGSPQPPKPYCSLRCPNHECRNDQHO
CCVQVHIDRCLRCPHCECKVDHNGKCCVRSRHPPEVCSLCPHCEKFDHG
KKCVYKIHCEVDCDIDCGRPECKLRHOGSKCVASERPHPOHCKNKRCPHGECK
VDQHEKCECVAAHRRPPKCSLRCPHRCRVNHFQEBCCVQVHDKCSLRCPHGECK
LHCAKTVPIDKENCCEKVPVCSNPLDGHGPT"
19762..21336
/note="Geneid exon scores (in order of location ranges) :
106.68 - GSCU_ID dd_00725"
/codon_start=1
/product="hypothetical protein"
/protein_id="AA338632.1"
/db_xref="GI:42733688"
/translation="MDNNNLQPPYFPGSNYTCDCDNDENGNDSTLTIOPKPAPBLPS
SQSPSPTIISQJNHAKYQPKFKEHNNNIDSKMLCVQISEQISPKATNTFSDMLN
SLMLPPTPFKLVNDHLKRYIIPFSYVNVNVEFSGYTDSKLNNTALYDKYNTN
STYNNNNYNNVNNNSNNSNADAKKKTITSATILKENVELYNNGNNNLNKNS
EILDKLCKIKNNKVGQYTRITNRLILACNSMDLSESSINNCVNSOOFDPNNYIEK
SLDENQOAOLODOVEYLAARYEDVYKWRDYLNIHLETOSMSIYKLOQESKFPYV
SSTIQPNKLSLVYIPTYTYTKEQNSFSPFLNSQNGVISAIRPEIGLCTGPFK
TYNNFSLTVNGFEPSKCDGLAIIDSPKYNTSSYFLTFRSSTYPLGSSSIGHL
TIIRNSKSIATLTLRQKRSRVSQKNPFLPHERTFSPKGMCFEIIIDGEVNOIDVI
KINPEGNINYSIGN"
complement(21961..24357)
/note="Geneid exon scores (in order of location ranges) :
173.48 - GSCU_ID dd_00724"
/codon_start=1
/product="similar to Arabidopsis thaliana (mouse-ear
cress). Putative PREG-1-like negative regulator"
/protein_id="AA338633.1"
/db_xref="GI:42733689"
/translation="MSKYSKQYVPSKPKPQNIPTALPSSVASLSSLOQPRNAG
SHRKQRGSSSKNPNPLDQISTEKSTSGNSKNNNNNNNNKRRKTSKNSNVNSNNIN
SSSIVIKCEBQ000000000000000000000000DGFLLKSGEPNLPFKNLYLKXDIS

SPSSSSSTSPSCISSSSSSTFSOSIEAENKODINNNCNTNGNFEEEREKRV
E00SIEERTIESFEETKEVREPEREOPELLIPISIGETIKETYSPLSYSELI
DMYNDKSSSTSSQSSSSSSSSSPSPKPKIKNNNNNTTTTITTTTTSKTATSS

Query Match 5.6%; Score 73.4; DB 3; Length 25769;

Best Local Similarity 49.4%; Pred. No. 0.00015;

Matches 217; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

```
QY 438 CCGAACAAGATTAAATAAACAAGTAAGTAACTAATCCATCGAAAGCTATCATGTA 497
DB 2392 CGAACAACACGCTAATAAATAAATAAATTTGTGAGAAAAAATAAATAAATAA 2451
QY 498 TGTAAAGAAACATCTATTAAACACGATCCCTTAATAAACAAGCATATTTGCAAG 557
DB 2452 TATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2511
QY 558 AGACAAATTATGTTACAGTTTACAAACATCTAAGCGCAAAATTATGAAAGTAG 617
DB 2512 AAATTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2571
QY 618 CTATGACGTTCAATTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 677
DB 2572 ATAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2630
QY 678 TTTCTCTCTTACAAATAGATGATTTTCTCCGATTTTATTAATAAGTAAAGTAT 737
DB 2631 TAATTTTAAAGATATTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2690
QY 738 TTTATTAAGACAGCATGCTAGATTCCTGTCAAAATCTTTCTGATTTTTTAAAG 797
DB 2691 TTAATAATTTGGTGTGATTAATAATATCTTTTAAATTTTTCGAAATATCTTT 2750
QY 798 AGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAATTAAGT 857
DB 2751 TTTTCTTTTCAACACCAACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2810
QY 858 TATTTCTCTTAAAT 876
DB 2811 TTAATTTTATTTTAT 2829
```

RESULT 10
AF034389/c
LOCUS AF034389 713 bp DNA linear INV 04-FEB-1999
DEFINITION Plasmodium falciparum sexual stage antigen (s16) gene, promoter and
partial cde.
ACCESSION AF034389
VERSION AF034389.1 GI:3098290
KEYWORDS
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1 (bases 1 to 713)
AUTHORS Decherling,K.J., Kaan,A.M., Mbacham,W., Wirth,D.F., Eling,W.,
Koning,R.N. and Stunnenberg,H.G.
TITLE Isolation and functional characterization of two distinct
sexual-stage-specific promoters of the human malaria parasite
Plasmodium falciparum
JOURNAL Mol. Cell. Biol. 19 (2), 967-978 (1999)
MEDLINE 99108072
PUBMED 9891033
REFERENCE 2 (bases 1 to 713)
AUTHORS Decherling,K.J., Kaan,A.M. and Koning,R.N.H.
TITLE Direct Submision
JOURNAL Submitted (12-NOV-1997) Molecular Biology, University of Nijmegen,
Toernooiveld 1, Nijmegen 6525 ED, The Netherlands
FEATURES
source 1..713
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/strain="NF54"
/db_xref="taxon:5833"

gene

/gene="s16"
1..336
/gene="s16"
537..>713
/gene="s16"
537..710
/gene="s16"
711..>713
/gene="s16"
/note="Pfs16"
/codon_start=1
/product="sexual stage antigen"
/protein_id="AA012581.1"
/db_xref="GI:3098291"
/translation="M"

ORIGIN

Query Match 5.4%; Score 71.2; DB 3; Length 713;
Best Local Similarity 45.8%; Pred. No. 0.00045;
Matches 282; Conservative 0; Mismatches 333; Indels 1; Gaps 1;

```
QY 444 CAAGATTAAAAAATAACCAAGTAATCACTCAAGAGCTATCATGTAATGTTA 503
DB 681 CAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622
QY 504 AAGAAACATCTATTAAACACGATCCCTTAATAAACAAGCATATTTGAAAGACAA 563
DB 621 AATATATCTATGATAATTAAGCTATTTGTGAAATTAATTAATTAATTAATTAAT 562
QY 564 ATATGTACAGTTTCAACAACTATTAAGACGCAAAATTATTCGAAAGTAAAGTATGA 623
DB 561 ATATATGTAATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 503
QY 624 CGTTCAGATTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 683
DB 502 TTTTCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 443
QY 684 CTCTTCAATTAAGCATTTTCTCCGATTTTATTAATAAGCATTAAGTCAATTTTAA 743
DB 442 TTTTCAGTAAATAAGTATTAACGAAATATGTAAGAAAGAAATTTGGTAG 383
QY 744 TAAGACACGACATGCTGATTCGTCAAAATCTTTGATTTTAAAGCTAG 803
DB 382 TAAATTTTATTAATAATTAATAAAGAAATTAATTTTATTAATAATTAATAAAG 323
QY 804 TTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTATTTT 863
DB 322 TTTATTTATTTATTAATTAATTAATTTTAAATTTTAAATTCATATTTTAAATTT 263
QY 864 CTCTTATTAATAAGAAACCTTGAATAAATAGAGTTGCCAGACTAGCCCTGAAGTTT 923
DB 262 AATATTTTATTTTATTTTATTTATTAATAATTAATAATTAATAATTAATAATTT 203
QY 924 TTCCCAATAATTAACATCATCTGTAATTAATTTGGCAGCCCATTAATTAATTA 983
DB 202 TTTTCTTTTATTTTATTTATTTGAATTAATTTATTTTCTACATTAATAATAA 143
QY 984 CCGAACTGAATTCGACGAAACCAATCTGACTATTTCTGATTAAGTAAAGGGA 1043
DB 142 TAAATTTATTAATTTAAGGTATCATTAATGTTTTTTTAAATAAATAAATAATAGT 83
QY 1044 GAGAGAGAGAAAGAA 1059
DB 82 TATCTATATTAATAA 67
```

RESULT 11
DD086962/c
LOCUS DD086962 3576 bp DNA linear INV 23-JUL-2001
DEFINITION Dicyostelium discoideum Miga (miga) gene, complete cds.
ACCESSION U86962
VERSION U86962.1 GI:1841871
KEYWORDS


```

UNRES      location/Qualifiers
source      1..3576
            /organism="Dictyostelium discoideum"
            /mol_type="genomic DNA"
            /strain="AX4"
            /sub_strain="AK244"
            /db_xref="taxon:4669"
            /chromosome="4"
            /map="4.25"
            884..3423
            /gene="migA"
            join(<884..2899,2998..3423)
            rRNA
            gene

```

ORIGIN /number=2

RESULT 12
BX957346/c

VERSION BX957346.13 GI:54019944
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE	ORGANISM
Danio rerio (zebrafish)	Danio rerio

REFERENCE
1 (bases 1 to 149526)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

AUTHORS Beasley, H.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-hel@anger.ac.uk
Clome requests: clomerequests@anger.ac.uk
On Oct 9, 2004 this sequence version replaced g1:53850295.

```

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zlish-help@sanger.ac.uk
Project Information
Center project name: ZC117K16
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149526 bases at least Q40
Consensus quality: 149526 bases at least Q30
Consensus quality: 149526 bases at least Q20
Insert size: 149526; sum-of-contigs

```


EKGLDAEBAKEIKHLRQMLEQAGVRDLAAVGPCTEGVAEQNTIMDKFLDEBELKEAE
OCKNCPCPAQOQGPGBARSADSPRAGTGDHPADPADDDEDDDDDDDDDDDEEBEED
PCKKYNPDLISDDPTKQYGDCHERKNGKNGPKGKDIITVDDTVCMRPEERREBCL
YTIHAESETKNITODDLDAFIRTAALFTPSWQYTKIKNGADAQOLNNGITPEBEL
RSNYFTYGYDRICLNTDISKTVNDVAKAKDIKGFPSKDGSSPSGTTTPDDMWQTY
GDIIMWGMICALHVTNTEKTKIKINDSYDNQSONGNSPLEDPFAKPPFLRMKI
ENGEERCAERKLEONIGKSCNGINPIQYCSNDRHPCKNACDEKNYVETKKEFPGQ
TKFEVADLENADQEKDYKTQGPSKQGNIDKDKCNKCSQMEGAVLIDVSSKK
PGIYVAHKYSEKNCGAKGAPVTNVPAPRPPPPAPAPATTGGVAPCIVLTPF
SDNKEKCACTLKCGPKAPVTSMKCIPTGNTNMGAAITDSEGDASRHRDAPASGS
NQSICVPPRRKLIYPLTKMAEETTKSGSQESKAGTSSSSSSASPSGTSISQ
GERSFGSLSTPASTSSPSNRDDLLKAEVESAAYETFLMKYKMDKKELEDEKQ
OBSKGLVGLDNGSNVDDKDPQKLEKGIPEEFKQMFYTLGDYDIDILVRGNT
SUGNTNGSNNNNIIVIEASGDKODEMKIKOKAIDENISLKOAAVSPNRGOQOON
SSLTRETLMKEHAPSIWEGMICALTYKEMDEKTIYDNVEYREGTGTGTSKYKE
KXEYNTVUKLSDNSDTEAKTKAPASDNPFTLSHVLPAPPRYIIEBEGTFCBERK
KLAQIKVDCQSDADYKCSGGEBCIKEDISITGVFADLCGCGERCKRKYKWIER
KILFEQOKSAVYKQTKCKESGGGNGVCTVKCTDPAQFLERLGCNKNDNGEG
KLEFNEQOSTFKHTKHCDPCSFKLIDCRANGKCSGDTGKCDGIITTIIDAKELAKNIS
TPDVVRVSDNDNTPEGDDLKVCCEGKGFKGIKREMKCRNCGGLDVGCLKKGND
KLDDKOIILIRALIKRMLFLEEDYKIKHKISDCTNNGENICRDKCONKNCVEM
ITLKREMEKIKIKHYLEKKEGNDMKSSVRNPLEKEHREPERNAIKRCKLTQPS
FGLNDKPSQNGHODADICMIKQLDEKITSLSSTSGQETECQEHPLDDDEL
LBETENPVKQOPSPCPKVEKKEIYDEKCGEDTSEPKTDBESKXKXKDBSTDP
EBPPATPEAPKEKKEVBPQOPQPTTPOIVDKTPTALVTSTLAWSGIGFAATYFEL
KKTKSTIIDLAVINIPKSDYDIPKLSPNRYIPTSQYRGKRYIYLEGDSGTSGY
THYSDITSSSESEYEMDINDIYAPAPRYKTLIEVLSEPSGNNTTASGNNTTASGN
NTTASGNNTTASGNNTPSDNDIQSDGIPSSKITDNEMNTLKDEFISQYIOSQPKD
VNDTSSGDIPTFOTONTLYFDKPEKPTISIHBNKLYTGEYNTDSTNGNDLY
NGKNNLVSGQNNVYSGIDPTSDNRGLTSGHDSYSGIDILINJLSGNQHIDILYDEK
KRENELEFNTNHYHTTINRFAPARDPILNOLLEFHTWLDHRNCKEMNKKEBLD
KKEWENETHSGNTHPSDNTKLTANDPILQIDMDPKPINQFTNMDIVDPTNMDM
EDDIYDVNDHDSITVDSNTMDVPSKVQIEMDVNTKLVEKPIADVMDI"
45458 46601
/gene="PF00010w"
/gene="PF00010w"
45458 .46601
/gene="PF00010w"
/note="VAR exon II, could be alternate exon II for
upstream VAR gene PF00005w"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="synonym: PF00015c"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="1 probable transmembrane helix predicted by
TMHMM2.0 at aa 328-350
Similar to Plasmidium falciparum rifin pf00030C
SMALL:096112 (EMBL:AE001367) (370 aa) fasta scores: E():
3.4e-65, 53.6% id in 375 aa"
/codon_start=1
/product="RifIN"
/protein_id="CAD49095.1"
/db_xref="GI:23498128"
/db_xref="UniProt/TrEMBL:O81219"
/tranlation="MKIHYINILFELPILNITLQYQRNHSNLTHTONNRSCECELY
AATYDDDDPQMKVWVFKSQTOQRFHEDEDMVEYKQCKCKEIKLILKXME
KELMDKFLATIDTIDSDAIPTCVCEGLADKVEKGLGKGVGTAVPVLGIGAV
HYMKPALKALAIKAIAGSTADIDAAEAAGAKRGMEFYIKLKGTVENPFGICDT
ISTSTGYTVTEFVNTIYKNGTCHUMSSINPTACTYITELSLKTKGATGCHPP
LAIIRQMIKGLABEATEAKAABAANAQALTAITKEQYALIEAGNSSITSINASTI
ALITIVLWVIYIILIRYRRKKMKKKYIKLKE"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="synonym: PF00202c"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="Similar to Plasmidium falciparum variant-specific
surface protein var-3 SMALL:026032 (EMBL:L40609) (3306 aa)
fasta scores: E(): 6.2e-103, 43.31% id in 3551 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PEMPL)"
/protein_id="CAD49096.1"
/db_xref="GI:23498129"

Qy	444	CAAGATTTAAAAAATTAACAAGTACTGATCCAGTGAAGTACTGATGTTTAA	503	/db_xref="GOA:O81218"
Qy	335664	CAAAAAGAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA	335605	/db_xref="UniProt/TrEMBL:O81218"
Qy	504	AAGAAACATCTATTAACCAACGATCTTTAAAAAACAAGCATATTTGAAAGACA	563	/translation="WGTSSTPSPVSKDVANBSHNAVLENIGIEIYENBEKKNY
Qy	335604	AATATATTCCTATGATTAATAAGCTAATGGAATTAATTTATGAAAAATATA	335545	/trnslatlon="WGTSSTPSPVSKDVANBSHNAVLENIGIEIYENBEKKNY
Qy	564	ATTATGTTACGTTTACAAACATCTTAAGACGACAAATTATATGAAAAGTAA	623	NRFEENEAACNSDKIRVTGESSKAGGACAPPRNKNLKYIDNTDITDIDL
Qy	335544	ATATATGTAATATATTTTATTAATATATATATATATATATATATATATAT	335485	NVLVTAKENSVAKHAKENSVAKHAKENSVAKHAKENSVAKHAKENSVAKH
Qy	624	CGTTCAGATTTTCTTTTCAATCTGTTATTTGTTATTTGTTATATACATTTCT	683	RAVFETKIDNLEKEIISDNDPNYVTKLEEDMTANRQWMAITCYIPLYNYFEK
Qy	335484	TTTTTTTTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	335425	SDDTLVFNNDKEKGEYEGAPPTNIDYVOPPLRWPEMBAEPRINHHLRKIKD
Qy	684	CTCTTACAATAGAGTATTTCTTCGATTTTATATAAATGACTATTAAGTATTT	743	DKREIYVNGONQCTKRIEKGSSCKREKCSKCYVDYDIEKQNEFQDQYK

Query Match 5.4%; Score 70.4; DB 3; Length 347582;
Best Local Similarity 44.6%; Pred. No. 0.00055;
Matches 275; Conservative 0; Mismatches 341; Indels 0; Gaps 0;


```

Db      335424  TTTTCAGTATATTAAGCAATATGAGAAATGGAAGAAAGAAATTTGGATG 335365
Qy      744  TAAGAGACCGAGTCGTAGATTCCTGTCGTAATTTTTCGATTTTTCAGAGCTAG 803
Db      335364  TAAATTTTAAATATATAATATAAAGAAATATATTTTATATATATATATAAAG 335305
Qy      804  TTTGGCAACCTGTTCTTCAAGAAATTTTGATTTTTCAGAAAAATTAAGTTATTT 863
Db      335304  TTTATTTATTTTATATATATATATATTTTAAATTTTAAATTCATATTTTAAATTT 335245
Qy      864  CTCTTTATTAATAATGAAAAACCTAGAAAAATAGAGTTGCCAGACTGACCTAGATGTT 923
Db      335244  ATATTTTATTTTATTTTTCCTTATATATATATATATATATATATATATATATTT 335185
Qy      924  TTCCCAATTAATTAACAATCACTGTATTAATTTTGGCAGCCCATTAATTTTAA 983
Db      335184  TTTTATTTTATTTTATTTTGAATATATTTATTTATTTTCTACATATATATAATA 335125
Qy      984  CCGAAATCGAAATCGAGCGAAACCAATTCGAGCTATTTCTTATGATTAAGTAAGGGA 1043
Db      335124  TAAATTTTATTTTATTTTAAAGTGATCAATTAATGTTTTTAAATTAATAAATATGCT 335065
Qy      1044  GAGAGAGAGGAGAGAA 1059
Db      335064  TATCTATATATATATAA 335049

```

RESULT 14
CQ422641/c 883 bp DNA 1linear PAT 28-JAN-2004

```

LOCUS      CQ422641
DEFINITION Sequence 7675 from Patent WO0151628.
ACCESSION CQ422641
VERSION    CQ422641.1 GI:41374870
KEYWORDS
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1
AUTHORS    Lillie,J., Xu,Y., Wang,Y. and Steimann,K.
TITLE       Novel genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of breast cancer
JOURNAL     Patent: WO 0151628-A 7675 19-JUL-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES
SOURCE     1. 883
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN

```

Query Match 5.4%; Score 70.2; DB 6; Length 883;
Best Local Similarity 39.2%; Pred. No. 0.0007;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

```

Qy      419  TGTCCTACTATGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAAGTAATCCA 478
Db      873  TGTNNNNNNNACCTTAACCTTTAAATATNTNNNANTNNANNAAAAAAATATTTAANN 814
Qy      479  CTGGAAGCTATGATGATGTTTAAAGAAACATCTATTAACCAAGACCTCTTAA 538
Db      813  TAAAAATTTTNNAAAAAATTAATNTNANNATNNTNNTNNTNNTNNTNNTNNTNNT 754
Qy      539  AACAGCATATTTGAAAGAGACAAATATGATTAAGTTTCAACAACTTAAGAGCACA 598
Db      753  TTAANNNTTTTNTAANTATTAACCAAAATTTTAAAAAATTTTAAAAAANTT 694
Qy      599  AATTATATCGAAAGGTAGCTAGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTG 658
Db      693  AATTAAATTTAAATTTTATTAATNAAAAAATTTTAAATTTTAAACAAANTTTT 634
Qy      659  TTATTTGTTTATATACATTTTCTCTTCAATAGAGATTTTCTTCGATTTTATA 718

```

```

Db      633  TTTNTTNNATTAATAAATTTTAAATTAATAAANNANNTTTTATATNANATTTAA 574
Qy      719  AAATGACTATPAAGTATTTTATATPAAGACGAGATGCTAGATTCCTGTCAAAAA 778
Db      573  AAAAAAANNTTTTATTTTATATNTTATTAATAAATAATTTATTTNTTTTTCNAAAANA 514
Qy      779  TCTTTCTGATTTTATTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
Db      513  AAAAAAATTTTNNAAANNTTTTAAACCTTNAATTAANANANATTTTNTTNTT 454
Qy      839  TTTTCAAAAAATTTAGTTTATTTTCCTTATTAATTAATAAGAAACCTTAGAAAAATAGA 898
Db      453  NNANATTAATAAATTTATTTTNTTCTATTAATAAATAAANANANATTAATAAT 394
Qy      899  GTTGCAAGTACCCCTAGATGTTTCCCAATTAATTAACAATCACTGTGTATATATTT 958
Db      393  TTTNAAAAATATATTTTATATNTATNTNANANANAAATTTAAATNANANAAAAATTTTA 334
Qy      959  TGCCAGCCCATTAATTTATTAACCGAAACTGAAATCGAGGAAACCAATCTGAGCT 1018
Db      333  NTTAAATTAATTAATAAATAATTTAAACNTCATTTNTAATTTATTAATTAATAAATAAATTT 274
Qy      1019  ATTT 1022
Db      273  NTTT 270

```

RESULT 15
BV119878/c 241 bp DNA 1linear STS 19-MAR-2004

```

LOCUS      BV119878
DEFINITION PZA01377 CML247 Zea mays CML247 Zea mays STS genomic, sequence
            tagged site.
ACCESSION BV119878
VERSION    BV119878.1 GI:45589251
KEYWORDS
SOURCE     Zea mays
            Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 241)
JOURNAL     McMullen,M.D., Vroh Bi,I., Schroeder,S.S. and Gaut,B.S.
            MPZ-OCI Joint SNP Discovery
            unpublished (2003)
COMMENT

```

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTCTACGTCGAATGATGATCAAGA
Primer B: AGGAAACCAAGCTTATGAGAGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: Redtaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with ampliTaq DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
Redtaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

PHRED/PHRAP Quality Scores 56 56 56 51 51 40 40 40 35 35 42 48
56 56 56 56 42 42 56 33 33 33 47 44 37 37 37 37 48 48
42 42 42 30 30 30 37 37 44 44 44 84 84 90 89 81 71 68 51 40 45
45 51 39 39 56 44 48 86 86 90 81 71 68 63 62 62 58 58 69 69
67 67 77 72 68 71 68 72 74 80 74 76 73 66 77 79 68 62 57 56 53 60
58 71 71 68 68 77 75 79 71 79 68 81 75 90 79 35 37 33 39 42 42 35
37 30 38 20 48 40 44 40 33 42 37 37 74 68 60 50 54 54 51 51 55
49 50 57 72 82 90 90 88 88 79 74 74 67 64 62 53 50 52 59 55
65 71 71 84 84 84 88 85 90 84 88 88 90 80 80 79 83 83 86 86 81
76 68 49 48 39 39 42 49 33 42 29 24 24 14 14 32 44 40 48 48 90
84 84 84 87 79 72 72 63 57 53 51 41 57 54 56 56 49 51 30 33 33 28
28 28 33 40 42 48 37 37.

FEATURES

source

1..241

Location/Qualifiers

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="CML247"

/db_xref="taxon:4577"

/clone_lib="Zea mays CML247"

/dev stage="seedling"

/note="Organ: leaf; genomic DNA from inbred line"

STS

ORIGIN

Query Match

Best Local Similarity 61.5%; Score 70; DB 11; Length 241;

Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
QY      799 GCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATTAGTTT 858
         |||||
Db      195 GTTAGTTGGAAACCACTTTTCCCAAGGAATTCATTTTCCCAAGAAATTAGTTC 136

QY      859 ATTTCTCTTTAAATAAGAAACACTAGAAATAAGTAGTCCAGACTAGCCCTAGA 918
         |||||
Db      135 ATTTCCCTTGGAAATAGAAATCCCAAGGAAATGAGTCCCAACTAACCTTAAT 76

QY      919 ATGTTTCCCAATTAATTAACAATCACTGTGTATTAATTATTGCGCAGCCCATTAATAT 978
         |||||
Db      75  ACTGTTAATTATTGTCATATCCATCACCAGAGATGATATCTTAACCGCATATCTTGT 16

QY      979 TT 980
         ||
Db      15  TT 14
```

Search completed: September 16, 2005, 03:01:14
Job time : 4402.14 secs


```

Qy 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGAGGTC 60
Db 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGAGGTC 60
Qy 61 CTTCTTCTGCTTATTACTGATGAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTACTGATGAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
Qy 121 CTGCAATGTTCACTCTTCCCACTCGGGTGGACATTTCTTGATGTCGGTGGTCCCAT 180
Db 121 CTGCAATGTTCACTCTTCCCACTCGGGTGGACATTTCTTGATGTCGGTGGTCCCAT 180
Qy 181 CTGACCGAGGSCCATGACACCTTTGGGACACCCATCAAGGGGCTTTGGATGGCCCA 240
Db 181 CTGACCGAGGSCCATGACACCTTTGGGACACCCATCAAGGGGCTTTGGATGGCCCA 240
Qy 241 CGAAGCGTATCGGTCGTGATGATCCAGGGGATATATGTCCTCCACAAATGTCACCTATA 300
Db 241 CGAAGCGTATCGGTCGTGATGATCCAGGGGATATATGTCCTCCACAAATGTCACCTATA 300
Qy 301 TTATATCTTCTTATGATTTTATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATCTTCTTATGATTTTATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGGCTCAGCATGATTTGCTTAGAGGCCAGAAATGAGGAGCCAGCATGCTTAGTG 420
Db 361 GGGGCTCAGCATGATTTGCTTAGAGGCCAGAAATGAGGAGCCAGCATGCTTAGTG 420
Qy 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Db 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Qy 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATGTAAGCGACAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATGTAAGCGACAA 600
Qy 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTCTGTAATTTGTT 660
Db 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTCTGTAATTTGTT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTCTTACATATGAGTATTTTCTTCGATTTTATAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTCTTACATATGAGTATTTTCTTCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATATTAAGAGACGCGATGCTGATGTCGTTCAAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATATTAAGAGACGCGATGCTGATGTCGTTCAAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATAATGAAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATAATGAAAACACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTGAAATGTTTCCCATTAATTAATGATCACTGATTAATTTTATG 960
Db 901 TGGCAGACTAGCCCTGAAATGTTTCCCATTAATTAATGATCACTGATTAATTTTATG 960
Qy 961 GCCAGGCCCATTAATTAATTAACCGAACTGAAATGACGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGGCCCATTAATTAATTAACCGAACTGAAATGACGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTATGATTTAGTAAAAAGGAGAGAGAGAAATCACTGTTTAAGTCAATGTCCTC 1080
Db 1021 TTCTCTATGATTTAGTAAAAAGGAGAGAGAGAAATCACTGTTTAAGTCAATGTCCTC 1080
Qy 1081 TGAGATGTGGGTTTGGCAACGATAGCCAGTAAATCATAGTACATAGGTGCTAGGTCA 1140

```

```

Db 1081 TGAGATGTGGGTTTGGCAACGATAGCCAGTAAATCATAGTACATAGGTGCTAGGTCA 1140
Qy 1141 GGTTCGGAGCTCTGTCATCTCATGATGACATGACATGCTTGTTCACCGTTGTC 1200
Db 1141 GGTTCGGAGCTCTGTCATCTCATGATGACATGACATGCTTGTTCACCGTTGTC 1200
Qy 1201 TTGTTTCATGTCGCCAGGCTTGGCTTATTTGTAACCAAGAGATACCTACCTCCAAACAT 1260
Db 1201 TTGTTTCATGTCGCCAGGCTTGGCTTATTTGTAACCAAGAGATACCTACCTCCAAACAT 1260
Qy 1261 CCATCTTACATGCAACTTCCATGCAACAGCAGACATATGTTCTGAAAC 1311
Db 1261 CCATCTTACATGCAACTTCCATGCAACAGCAGACATATGTTCTGAAAC 1311

RESULT 3
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; de.
XX
OS Zea mays.
XX
PN MO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGAGGTC 60
Db 1 CCAATGCTCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGAGGTC 60
Qy 61 CTTCTTCTGCTTATTACTGATGAAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120

```


Db 61 CTTCCTCTGCTTATTAAGTGAATCGGGGTTACAAAAAATTCACGCGTGCAATGAT 120
 QY 121 CTCATGTTTCACGTTCCGACCTCGGGTGCACATTTCTGGATGCGGTGGTCCCAT 180
 Db 121 CTCATGTTTCACGTTCCGACCTCGGGTGCACATTTCTGGATGCGGTGGTCCCAT 180
 QY 181 CTGACCGAGGCCCATGAGACACCTTTGGGACACCATCAAGGGCCCTTGGATGGCCCA 240
 Db 181 CTGACCGAGGCCCATGAGACACCTTTGGGACACCATCAAGGGCCCTTGGATGGCCCA 240
 QY 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAACATGTGCACTAT 300
 Db 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAACATGTGCACTAT 300
 QY 301 TTAATATCTTTAGATATTAATTTTGAATAAATAACAATTAATCTTTTGGTGA 360
 Db 301 TTAATATCTTTAGATATTAATTTTGAATAAATAACAATTAATCTTTTGGTGA 360
 QY 361 GGGCCTCAGCATGATTTTGCCTTAGGGCCAGAAAATGCGAGGACGAGCCATGTCTAGT 420
 Db 361 GGGCCTCAGCATGATTTTGCCTTAGGGCCAGAAAATGCGAGGACGAGCCATGTCTAGT 420
 QY 421 TCCACTATTTGGCACTACCCAGAAACAAGATTAAATAAATACCAAGTAATCATCACT 480
 Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTAAATAAATACCAAGTAATCATCACT 480
 QY 481 CGAAGCTATCATGATATGTTTAAAGAAACATCATTTAAACACGATCCCTTAAATA 540
 Db 481 CGAAGCTATCATGATATGTTTAAAGAAACATCATTTAAACACGATCCCTTAAATA 540
 QY 541 CAAGCATATTTGAAAAGAGCAAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
 Db 541 CAAGCATATTTGAAAAGAGCAAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
 QY 601 TTAATTCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTATCTTGTATTTTGT 660
 Db 601 TTAATTCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTATCTTGTATTTTGT 660
 QY 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTAA 720
 Db 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTAA 720
 QY 721 ATGACTATTAAGTCAATTTTAAAGAGCAGCATGTGTAGATTTCTGTTCAAAAATC 780
 Db 721 ATGACTATTAAGTCAATTTTAAAGAGCAGCATGTGTAGATTTCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGTATTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
 Db 781 TTTCTGATTTTTTAAAGAGTATTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
 QY 841 TTTCAAAAAAATAGTTTATTTTCTTCTTAAATAGAAAACACTTAGAAAATAGAGT 900
 Db 841 TTTCAAAAAAATAGTTTATTTTCTTCTTAAATAGAAAACACTTAGAAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTCAATCACTGTATTAATTTT 960
 Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTCAATCACTGTATTAATTTT 960
 QY 961 GCGAGGCCATTAATTTTAAACCGAAAATCGAAATCGAGGAAACCAATCTGACTAT 1020
 Db 961 GCGAGGCCATTAATTTTAAACCGAAAATCGAAATCGAGGAAACCAATCTGACTAT 1020
 QY 1021 TTTCTAGATTAATAAAG 1080
 Db 1021 TTTCTAGATTAATAAAG 1080
 QY 1081 TGAAGATGTGGGTTTGGCAACGATAGCCACCGTAATCATAGTCTATAGTGCTACGTCA 1140
 Db 1081 TGAAGATGTGGGTTTGGCAACGATAGCCACCGTAATCATAGTCTATAGTGCTACGTCA 1140
 QY 1141 GGTTCGGGAGCTCTCGTGCATCTCAATGGCAATGCTTGTCTTCAACGTTTCGTC 1200
 Db 1141 GGTTCGGGAGCTCTCGTGCATCTCAATGGCAATGCTTGTCTTCAACGTTTCGTC 1200

QY 1201 TTTGTTCCATGCTGCAAGCCTTGCTTATTTGAACCAAGAGATPACTACTCCAAACAT 1260
 Db 1201 TTTGTTCCATGCTGCAAGCCTTGCTTATTTGAACCAAGAGATPACTACTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311
 Db 1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 4

AAH76333 standard; DNA; 1394 BP.

AAH76333;

29-OCT-2001 (first entry)

Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

Ms45; male tissue; regulatory region; transcription; male fertility; hybrid seed; de.

Zea mays.

MO200160997-A2.

23-AUG-2001.

13-FEB-2001; 2001WO-US004527.

15-FEB-2000; 2000US-00504487.

(PION-) PIONEER HI-BRED INT INC.

Albertain MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

WPI; 2001-514772/56.

PT A male tissue-preferred regulatory region comprising nucleotide sequences essential for initiating transcription of the MS45 gene useful for mediating fertility in a male plant.

Claim 4; Page 47; 50bp; English.

CC The invention provides a male tissue-preferred regulatory region (I) comprising nucleotide sequences essential for initiating transcription of the MS45 gene. A method of mediating male fertility in a plant is provided that involves introducing an expression vector comprising a promoter operably linked to (I) into a plant where the exogenous gene impacts male fertility of the plant and (II) controls expression of the exogenous gene. A method of producing hybrid seeds is also provided. The present sequence represents a nucleic acid sequence encoding an Ms45 male tissue preferred regulatory region from Z. mays

Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 100.0%; Score 1311; DB 5; Length 1394;

Best Local Similarity 100.0%; Pred. No. 1e-269; Mismatches 0; Indels 0; Gaps 0;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCATGATGCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGGGTCC 60

1 CCATGATGCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGGGTCC 60

61 CTTCCTTGCCTTATTAAGTGAATGGGGTTACAAAAAATCTTCACGGGTGATGAT 120

61 CTTCCTTGCCTTATTAAGTGAATGGGGTTACAAAAAATCTTCACGGGTGATGAT 120

121 CTCATGTTTCCACTCTCCACCTGCGTTGACATTTCTTGGATGGGTGTTCCCAT 180

121 CTCATGTTTCCACTCTCCACCTGCGTTGACATTTCTTGGATGGGTGTTCCCAT 180

QY 181 CTGACGAGGCGCCATGACACCTTTGGGACACCATCAAGGGCTTTGGATGGCCCA 240
 DB 181 CTGACGAGGCGCCATGACACCTTTGGGACACCATCAAGGGCTTTGGATGGCCCA 240
 QY 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCCAACATGTCCTATA 300
 DB 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCCAACATGTCCTATA 300
 QY 301 TTATATCTTTAGATATATTTAATTTTGGAAAAATTAACAACTTATCTTTGTGA 360
 DB 301 TTATATCTTTAGATATATTTAATTTTGGAAAAATTAACAACTTATCTTTGTGA 360
 QY 361 GGGCCCTCAGCATAGATTTTGGCTTAAAGGGCCAGAAATGCGAGACCATGCTAGTG 420
 DB 361 GGGCCCTCAGCATAGATTTTGGCTTAAAGGGCCAGAAATGCGAGACCATGCTAGTG 420
 QY 421 TCCACTATTTGGCACTTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
 DB 421 TCCACTATTTGGCACTTACCCAGAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
 QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAAGTAACTATCCACT 540
 DB 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAAGTAACTATCCACT 540
 QY 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATGTAAGGAGCAAA 600
 DB 541 CAAGCATATTTGCAAGAGACAAATTTATGTTACAGTTTACAAACATGTAAGGAGCAAA 600
 QY 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCACTTCTGTATTTTGT 660
 DB 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCACTTCTGTATTTTGT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTCTTACATATGATGATTTTCTTCCGATTTTATAA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTCTTACATATGATGATTTTCTTCCGATTTTATAA 720
 QY 721 ATGACTATAAGTCAATTTTATATATAGAGCAGCATGCTGAGATTTCTGTTCAAAAATC 780
 DB 721 ATGACTATAAGTCAATTTTATATATAGAGCAGCATGCTGAGATTTCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTTCTTCAAGAAATTTTGAATTT 840
 DB 781 TTTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTTCTTCAAGAAATTTTGAATTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATTAAGAAACACTTAAAGAAATAGAGT 900
 DB 841 TTCAAAAAAATTAAGTTATTTTCTCTTATTAATTAAGAAACACTTAAAGAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCAATCACTGTATTAATTTATG 960
 DB 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCAATCACTGTATTAATTTATG 960
 QY 961 GCGAGCCCAATTAATTTAAACCGAAATCGAAATCGAGAAACCAATCTGAGCTAT 1020
 DB 961 GCGAGCCCAATTAATTTAAACCGAAATCGAAATCGAGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAAATCAAGTTTAAGTCAATGTC 1080
 DB 1021 TTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAAATCAAGTTTAAGTCAATGTC 1080
 QY 1081 TGAAGATGAGGATTTGGCAAGATAGCCAGATGATAGCTATAGCTATAGGCTCAGCTCA 1140
 DB 1081 TGAAGATGAGGATTTGGCAAGATAGCCAGATGATAGCTATAGGCTCAGCTCAGCTCA 1140
 QY 1141 GGTTCGAGACTCTCGTGTATCTCAATGAGCATATCAATGCTTGTCAACCGTTCTGTC 1200
 DB 1141 GGTTCGAGACTCTCGTGTATCTCAATGAGCATATCAATGCTTGTCAACCGTTCTGTC 1200
 QY 1201 TTGTTTCATGCTCGAAGCTTTGCTTATTTGAACCAAGAGATACCTATCCCAACAT 1260
 DB 1201 TTGTTTCATGCTCGAAGCTTTGCTTATTTGAACCAAGAGATACCTATCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCATGCAAAACGAGCATATGTTTCTGTAAC 1311

DB 1261 CCATCTTACTCATGCAACTTCATGCAAAACGAGCATATGTTTCTGTAAC 1311
 RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 XX AAH76340;
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 promoter fragment.
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; promoter; db.
 XX
 OS Zea mays.
 PN W0200160997-A2.
 XX
 XX 23-AUG-2001.
 PD
 XX
 PF 13-FEB-2001; 2001WO-US04527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 XX Albertsen WC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 PI WPI; 2001-514772/56.
 DR
 XX
 XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 XX Example 5; Fig 8; 50pp; English.
 PS
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
 Query Match 12.3%; Score 160.6; DB 5; Length 255;
 Best Local Similarity 97.6%; Pred. No. 1.3e-24;
 Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1145 CGGAGACTCTCGTGTATCTCAATGAGCATATGCTGTTCAACGTTGCTGTAT 1204
 DB 5 CGGAGATCCGCTGATATCATGACATGCAATCACTATGCTGTTCAACGTTGCTGTAT 64
 QY 1205 TCCATGCTCAAGCTTGGCTTATTTGAACCAAGAGATACCTATCCCAACATTCAT 1264
 DB 65 TCCATGCTCAAGCTTGGCTTATTTGAACCAAGAGATACCTATCCCAACATTCAT 124
 QY 1265 CTACTCATGCAACTTCATGCAAAACGAGCATATGTTTCTGTAAC 1311
 DB 125 CTACTCATGCAACTTCATGCAAAACGAGCATATGTTTCTGTAAC 171
 RESULT 6
 AAH76334
 ID AAH76334 standard; DNA; 158 BP.
 XX
 XX AAH76334;
 AC
 XX

DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; de.
OS
XX Zea mays.
XX WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen WC, Fox TW, Garnat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
Query Match 11.1%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 1.5e-21;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTGTCACTTCACATGGCATCTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC 1213
DB 1 CGTGTCACTTCACATGGCATCTACATGCTTGTTCACCGTTCGCTTGTTCATCGTC 60
QY 1214 CAAGCCTGCTTTCGTAAGCAAGAGATACCTACCTCCCAACATCCATCTTACTCAT 1273
DB 61 CAAGCCTGCTTTCGTAAGCAAGAGATACCTACCTACCTCCCAACATCCATCTTACTCAT 120
QY 1274 GCAACTTTCATGCAAAACAGCAGCATATGTTTCTCGAAC 1311
DB 121 GCAACTTTCATGCAAAACAGCAGCATATGTTTCTCGAAC 158
RESULT 7
ID AAL15210/c
XX AAL15210 standard; cDNA; 883 BP.
XX
AC AAL15210;
XX
DT 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 7667.
XX DE Human breast cancer; cell marker; cyclostatic; BR.
XX KM Human; breast cancer; cell marker; cyclostatic; BR.
XX OS Homo sapiens.
XX

PN WO200151628-A2.
XX
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Little J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
PT
XX
XX Claim 1; Page 1378; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.4e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCACATATTTGGCACTACCCGAAACAGATTATAAATAACCAAGTAATCA 478
DB 873 TGTNNANNANACCTNACCTTTAAANATNTNNANNNTNAAANNAATAATTTAAANN 814
QY 479 CTCGAAGATCACTATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAA 538
DB 813 TAAATAATTTTAAATACTAATTTTAAANNAATNNAATTTTNTTAAANNAANNA 754
QY 539 AACAGCATATTTGCAAGAGACAATAATATGTACAGTTTACAAACATCTAAGAGGACA 598
DB 753 TTAANNNTTTTAAANTATTTAAACCAAAATTTTAAAAAATTTTAAAAAANTT 694
QY 599 AATTAATTCGAAAGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTATTTTG 658
DB 693 AATAANTTAAATTTTAAATNNAATAAATAAATAAATAAATTTTAAANCAANTTTTTT 634
QY 658 TTAATGTTTATATACATTTCTCTCTCTTAATAAGAGTATTTCTTCGATTTTANA 718
DB 633 TTTNTTNNATTAATAAATAAANTTTTAAATTAATAAANNAANTTTTATTAANATTTTA 574
QY 719 AATGACTATAAAGCATTTTATATTAAGAGACGATGCTAGATTCCTGTCAAAAA 778
DB 573 AAAAAAANTTTTAAAAANTTTTAAANTTTTAAATAAATAAATTTTATNTTTTCTNAAAA 514
QY 779 TCTTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
DB 513 AAAAAAATTTTAAANNTTTTAAACCTTAAATTAANNAANNAATTTTNTNTT 454
QY 839 TTTTCAAAAAAATAGTTTATTTCTCTTTAATAAATAGAAACACTTAGAAAAATAGA 898
DB 453 NNANATTAATAAATTTATTTTAAANTTTTAAATTAATAAANNAANNAATTAAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAATCAATCACTGTGATTAATT 958

Db 393 TTTNAAAAAATTAATTTATATNTATNTNNAANAATAATTTAAATNTNANNAANAATAATTTTA 334
 QY 959 TGGCCAGCCCCATTAATTAATTTAAACCGAAATCGAAATCGAAGCAACCAATCTGAGCT 1018
 Db 333 NTTTAAATTAATAAATAATTAATTTAAACNTCATTTNTTAATTTTAATTAATAAATAAATAAT 274
 QY 1019 ATTT 1022
 Db 273 NTTT 270

RESULT 8

ACN85231/C
 ID ACN85231 standard; DNA; 960 BP.
 XX
 AC ACN85231;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Breast cancer related marker, seq id 6381.
 XX
 KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN US2003099974-A1.
 XX
 PD 29-MAY-2003.
 XX
 PF 18-JUL-2002; 2002US-00198846.
 XX
 PR 18-JUL-2001; 2001US-0306220P.
 XX
 PA (MIL-) MILLENNIUM PHARM INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 DR WPI; 2003-787014/74.
 XX
 PT Novel isolated polypeptide associated with breast cancer, useful for
 PT detecting presence of polypeptide in sample, as a marker for breast
 PT cancer.
 XX
 PS Disclosure, SEQ ID NO 6381; 36bp; English.

XX The invention relates to an isolated polypeptide (I) associated with
 CC breast cancer which is encoded by a nucleic acid molecule comprising a
 CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
 CC the polypeptide of the invention. The activity of the polypeptide of the
 CC invention may be described as cytostatic. The antibody is useful for
 CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
 CC invention are useful in the detection of breast tumours. (I) is useful as
 CC a marker for breast cancer and in breast cancer therapy. Sequences given
 CC in records ACN78851-ACN92934 represent nucleic acid markers associated
 CC with breast cancer. Note: The sequence listing does not form part of the
 CC specification but may be obtained in electronic format from the USPTO web
 CC site at seqdata.uspto.gov/sequence.html?docid=20030099974
 CC
 XX

Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
 Query Match 5.4%; Score 70.2; DB 11; Length 960;
 Best Local Similarity 39.2%; Pred. No. 3.5e-05;
 Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 419 TGTCCACTATTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAATCTAATCCA 478
 Db 933 TGTNNANNNNAACTTAACCTTTAAANATTTNNANNTNNAANNAANNAATATTTTAAAN 874
 QY 479 CTGAAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAACCAAGATCTCTTAAAA 538
 Db 873 TANAATAATTTTNAAAAACTAATTTNNAANNATNTNAATTTTNTTAAAAAANNNAAAAA 814

QY 539 AACAGCATATTTGGAAGAGCAATTAATGTTCAGTTTACAAACATCTAAGAGGACA 598
 Db 813 TTAANNNTTTTNTAANTTAATTAACCAAAATTTTTTTTTTAAAAAATTTTAAANNT 754
 QY 599 AATTATATCGAAAGTAAGCTATGAGTTTCAGATTTTCTTTTCATCTGTATTTTG 658
 Db 753 AATTAATTAATTTTNTTAATTAATAAATAAATAATTTTAAATTTTAAACAAATNTTTT 694
 QY 659 TTAATGTTTAAATACATTTCTCTCTTACAAATGAGTATTTCTCCGATTTTATA 718
 Db 693 TTTNTTTNNAATAAATAAATTTTAAATTAATAAATAAATAAATAAATAAATAATTTAA 634
 QY 719 AATGACTATTAAGTATTTTATTAAGAAGCAAGTGTGAGATCTCGTCAAAA 778
 Db 633 AAAAAAANNNTTTTTTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA 574
 QY 779 TCTTTCGATTTTAAAGCTAGTTTGCAACCTGTTTCTTCAAGAAATTTTGATT 838
 Db 573 AAAAAAATTTTNTAANNTTTTAAACCTTNAATTTAANNAATTTTNTNTT 514
 QY 839 TTTTCAAAAAAATTAATTTTCTCTTATTAATAAATAAATAAATAAATAAATAAATAA 898
 Db 513 NNAATTTAATAAATAATTAATTTTNTCTATTAATTAATAAATAAATAAATAAATAAAT 454
 QY 899 GTTGCAGACTAGCCCTAGAAATGTTCCCAATTAATTAATCACTGTATTAATTT 958
 Db 453 TTTNAAAAAATTAATTTTATTAATTAATTAATAAATAAATAAATAAATAAATAAATAA 394
 QY 959 TGGCCAGCCCCATTAATTTAATTAACCGAAATCGAAATCGAAGCAACCAATCTGAGCT 1018
 Db 393 NTTTAAATTAATAAATAATTAATTTAAACNTCATTTNTTAATTTTAATTAATAAATAAATAAT 334
 QY 1019 ATTT 1022
 Db 333 NTTT 330

RESULT 9

ADR04296
 ID ADR04296 standard; DNA; 13400 BP.
 XX
 AC ADR04296;
 XX

DT 04-NOV-2004 (first entry)
 XX

XX Corn FT homologue nucleotide sequence SEQ ID NO:63.
 XX

XX flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
 KW floral development; plant sterility; plant fertility; flowering time;
 KW plant growth rate; inflorescence architecture; tissue culture morphology;
 KW cell division; FT homologue; gene; ds.
 XX

OS Zea mays.
 XX

PN MO2004067723-A2.
 XX

PD 12-AUG-2004.
 XX

PF 29-JAN-2004; 2004WO-US002422.
 XX

PR 30-JUN-2003; 2003US-00343477.
 XX

PA (PION-) PIONEER HI-BRED INT INC.
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX

PI Danilevskaya O, Hermon P, Bruggemann E, Shitbroun D, Ananiev E;
 PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
 XX

DR WPI; 2004-580996/56.
 XX

PT New polynucleotides, specifically nucleic acid fragments encoding
 PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
 PT homologs, useful for floral development, e.g. engineering plant flowering

PT time.
PS Claim 6; SEQ ID NO 63; 109pp; English.
XX
XX The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (FTL), or Apetala3 (Ap3) homologue locus. Also
CC described: (1) a vector comprising the polynucleotide, (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;
Query Match 5.3%; Score 69.8; DB 13; Length 13400;
Best Local Similarity 73.6%; Pred. No. 8e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 811 ACCCTGTTCTTCAAGAATTTGATTTTCAAAAAAATAGTTATTTCCTTA 870
DB 8559 ACCATTATTATTTCAAGAGTTTATTATTTATCAAGAAAAATAGTTATTTCTCTGG 8618
QY 871 TAAATATGAAACATTTAGAAAAATAGATGCGACATGACCTTGAATGTTTCCCA 930
DB 8619 AAAAAATAAATTCATTAGAAAAATGGGTTTCAAACTAGTCTTATTAGTTTCCAT 8678
QY 931 T 931
DB 8679 T 8679
RESULT 10
ID AA210551 standard; DNA; 2657 BP.
XX
XX AA210551;
AC
AT
DT 16-NOV-1999 (first entry)
XX
XX DNA sequence of the P-Zip promoter of maize.
DE
XX P-Zip promoter; male sterile plant; glyphosate tolerance; glyphosate;
KW male reproductive tissue; hybrid seed production; crop outcrossing;
KM flower life; ss.
XX
XX Zea mays.
OS
PN WO946396-A2.
XX
XX 16-SEP-1999.
PD
XX 09-MAR-1999; 99WO-US005126.
PF
XX 09-MAR-1998; 98US-007277P.
PR
XX

PA (MONS) MONSANTO CO.
XX
XX PI Brown SM, Fromm ME;
XX
XX DR WPI; 1999-551420/46.
XX
XX PT Production of male sterile plants using a gene encoding glyphosate
PT tolerance, used for, e.g. production of hybrid seed.
XX
XX PS Disclosure; Fig 1A-B; 54pp; English.
XX
XX The present sequence represents the P-Zip promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC production of a second RNA sequence in a male reproductive tissue.
CC Expression of the DNA promotes tolerance to glyphosate in those tissues
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The
CC methods can be used with plants such as corn, wheat, rice, canola, oat,
CC barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
CC sunflower, soybean, tomato, cucumber and squash
XX
SQ Sequence 2657 BP; 686 A; 611 C; 677 G; 603 T; 0 U; 0 Other;
Query Match 5.1%; Score 67.2; DB 2; Length 2657;
Best Local Similarity 74.2%; Pred. No. 0.00019;
Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
QY 787 ATTTTAAAGAGCAGTGTGGCAACCGTTCCTTCAAGAATTTGATTTTCAAA 846
DB 1883 ATCTTTTAAGGCTAGTGTGGAAACACATT-TTCAAGGATTTCAATTTTGGCAG 1941
QY 847 AAAAAATGATTATTTCTTTATATAAATAGAAAAACCTTAGAAAAATAGATTGCCG 906
DB 1942 GGAATTAAGTTCATTTTCCCTTGGGAAATGAAATCCATGGAAAAATGTGTTCCCA 2001
QY 907 ACTAGCCCTAGA 918
DB 2002 ACTAGCCCTAAA 2013
RESULT 11
ID AAX58751 standard; cDNA; 6027 BP.
XX
XX AAX58751;
AC
AT
DT 16-AUG-1999 (first entry)
XX
XX Maize dull1 gene encoding starch synthase enzyme DUL.
DE
XX Starch synthase; SSII; DUL1 gene; maize; transgenic plant; ss.
KW
XX Zea mays.
OS
PN Key Location/Qualifiers
XX misc_feature 1..1437
FH
FT

FT /note= "functional fragment of cDNA specifically claimed
in Claim 16"
FT CDS 120..5147
FT /*cag= a
FT 120..1221
FT /*cag= b
FT /note= "functional fragment of cDNA specifically claimed
in Claim 12"
FT misc_feature 565..816
FT /*cag= d
FT /note= "functional fragment of cDNA specifically claimed
in Claim 14"
FT misc_feature 655..1221
FT /*cag= c
FT /note= "functional fragment of cDNA specifically claimed
in Claim 13"
FT misc_feature 1369..1944
FT /*cag= e
FT /note= "functional fragment of cDNA specifically claimed
in Claim 15"
FT misc_feature 1438..2424
FT /*cag= g
FT /note= "functional fragment of cDNA specifically claimed
in Claim 17"
FT misc_feature 2425..3791
FT /*cag= h
FT /note= "functional fragment of cDNA specifically claimed
in Claim 18"
FT MO924575-A1.
XX 20-MAY-1999.
XX 12-NOV-1998; 98WO-US024225.
XX 12-NOV-1997; 97US-00968542.
XX (IOWA) UNIV IOWA STATE RES FOUND INC.
XX Myers AM, James MG;
XX WPI; 1999-327406/27.
XX P-PSDB; AAY06199.
XX Nucleic acid encoding starch synthase enzyme of maize.
XX Claim 1; Page 104-107; 138pp; English.
XX This is the nucleotide sequence of the maize gene dnl1 (dnl). To
XX illustrate the role of the dnl locus in starch biosynthesis, a transposon
XX tagging strategy was used to isolate the gene and describe its
XX polypeptide product. The invention reports tagging of the dnl locus with
XX C₄ transposon, cloning and characterization of a portion of the gene, and
XX isolation of a near full-length cDNA (the present sequence). The amino acid
XX sequence (see AAY06199) deduced from this cDNA indicates that Dnl codes
XX for a 186 kDa protein extremely similar to potato tuber starch synthase
XX SSII1. Its expression pattern indicates that Dnl codes for SSII of maize
XX endosperm. The Dnl product contains unique sequence features in its N-
XX terminus that may mediate direct interactions with other starch
XX CC biosynthetic enzymes. Mutations within the maize SSII gene affect
XX multiple aspects of starch biosynthesis by disrupting an enzyme complex
XX containing starch synthase(s), starch branching enzyme(s) and possibly
XX starch debranching enzyme(s). The isolated cDNA can be used to provide an
XX enzyme with which to regulate the production of starch, and with which to
XX produce altered or novel forms of starch, e.g. in transgenic plants.
XX Expression of Dnl in bacteria and yeasts also modifies glycogen
XX production. Claimed expression vectors comprise the cDNA or fragments of
XX it that code for functional portions of Dnl
XX
SQ Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00035;

Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
Qy 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTCAAAAAA 851
Db 5606 TCTACGGGCTAGTTGGAAACCCCATTTT-TTCCAAGGATTTTCATTTTCCAAGAAAA 5548
Qy 852 TTAGTTATTTTCTCTTTATTAATAATGAAAAACCTTGAAAAATAGAGTTGCCAGACTAG 911
Db 5547 TTAGTTATTTTTCATTGAAAAAATTTGAATCTCTTGAAAAAATAGAGTTCACTACTAG 5488

RESULT 12
ABX0935/C
ID ABX0935 standard; DNA; 6027 BP.
XX
XX ABX0935;
XX
XX 17-FEB-2003 (first entry)
XX
XX DNA encoding maize Starch synthase III (Dnl).
XX
XX Starch; starch synthase; glucan association domain; GLASS; linker domain;
XX LINKR; glucosyl transferase domain; GLYTR; C-terminal end; CTEND;
XX granule bound starch synthase; GBS; morphology; retrogradation;
XX waterbinding; swelling potential; gene; db.
XX
XX Zea mays.
XX OS
XX
XX MO200279410-A2.
XX
XX 10-OCT-2002.
XX
XX 29-MAR-2002; 2002WO-US009574.
XX
XX 30-MAR-2001; 2001US-0279720P.
XX
XX (BADI) BASF PLANT SCI GMBH.
XX
XX Commuri P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;
XX WPI; 2003-040678/03.
XX
XX New DNA encoding fusion protein consisting of 4 different functional
XX PT domains selected from glucan association domain, linker domain, glucosyl
XX PT transferase domain, and C-terminal end, useful for producing modified
XX PT starches.
XX
XX
XX Claim 33; Page 225-227; 265pp; English.
XX
XX The invention describes an isolated DNA molecule encoding a fusion
XX CC protein consisting of 4 different functional domains selected from glucan
XX CC association domain (GLASS), linker domain (LINKR), glucosyl transferase
XX CC domain (GLYTR), and C-terminal end (CTEND) which are operably linked to
XX CC one another. The DNA molecule is useful for expressing in plants
XX CC polypeptides including starch synthase enzymes as fusion proteins with
XX CC improved affinity to starch and modified catalytic capabilities and to
XX CC the in vivo and in vitro synthesis of glucan chains of modified lengths
XX CC as compared to plants producing native starch or starch produced with
XX CC native starch synthases. Expression of the starch synthase fusion
XX CC protein along with granule bound starch synthase (GBSS) will lead to a
XX CC modified starch having an altered or improved morphology, retrogradation,
XX CC waterbinding, or swelling potential of the granules, gel strength,
XX CC adhesiveness, cohesiveness, hardness, elasticity, increased or decreased
XX CC granule size, degree of branching, crystallinity, degree of cross-
XX CC linking, and increased or decreased glucan chain lengths. This sequence
XX CC encodes a starch synthase used in the invention
XX
SQ Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

Query Match 5.1%; Score 66.4; DB 8; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00035;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Query Match 5.0%; Score 66.2; DB 8; Length 439;
 Best Local Similarity 54.3%; Pred. No. 0.0002;
 Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCATCTGTATTTGTATGTTTATATACATTTCTCT 685
 DB 182 TTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTTT 241
 QY 686 CTTACATAGAGTATTTCTCCGATTTTAAAGCATTAAGCATTTTATATA 745
 DB 242 TTTTATTT 301
 QY 746 AGACAGCAGATGCTGATTCGTCAAAATCTTTCGATTTTAAAGCTAGT 805
 DB 302 TGTCGGGGGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 361
 QY 806 TGGCAACCTGTTTCTTCAAGATTTGATTTTCAAAAAAATAGTTATTTCT 865
 DB 362 TTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATATCTTATTTTTTTTTT 421
 QY 866 CTTTATA 872
 DB 422 TTTCTTA 428

RESULT 15
 AA187279/c
 ID AA187279 standard; cDNA; 346 BP.
 XX
 AC AA187279;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 7339.
 XX
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukemia;
 KM nervous system disorders; arthritis; inflammation; se.
 XX
 OS Homo sapiens.
 OS
 PN WO200164835-A2.
 PN
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSB-) HYSBQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI, 2001-514838/56.
 DR P-PSDB; AAO07348.
 XX
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 XX and treating e.g. leukemia, inflammation and immune disorders.
 XX
 PS Claim 1; SEQ ID NO 7339; 1399bp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

QY Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;
 Best Local Similarity 52.6%; Pred. No. 0.00026;
 Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 446 AGATTTAAAAAATTAACCAAGTAACTATCCAGTGAAGCTATCATGTTTAA 505
 DB 334 ACAATTAAGTATATATTAATAATTAAGATTAATAATTTATCATTAATAATTTTA 275
 QY 506 GAAACATCTATTAAACCAAGATCTTTAAAAAACAAGCATTTTCGAAAGACAAAT 565
 DB 274 GACTATATATATATTAATAATTTTTTTTTAGAAAAATTAATTCATTAAACAATAAAA 215
 QY 566 TATGTTACGTTTACCAACCTTAAGAGCACAATTTATTCGAAAGTAAAGCTATGACG 625
 DB 214 TTTTAAATATTAATAATTAATAATTAATTAATTAATAATAATTAATTTTAT 155
 QY 626 TTCAGATTTTCTTTTCATCTGTATTTGTATTTGTTTATATATACATTTCTCT 685
 DB 154 TATTTATT 95
 QY 686 CTTACATAGAGTATTTCTCCGATTTTAT 717
 DB 94 TTTTTTTTTTTTTTTTTTTTTTTTTTCTTTTTTT 63

Search completed: September 15, 2005, 21:33:05
 Job time : 636.708 secs

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 198.304 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Perfect score: 1311
Sequence: 1 cccctggtcgtctccatgaaga.....cgacacatgcttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptcdatc/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptcdatc/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptcdatc/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptcdatc/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptcdatc/1/ina/PTCUS.COMB.seq: *
6: /cgn2_6/ptcdatc/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1311	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl1
2	1311	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl1
3	66.4	5.1	6027	2 US-08-968-542C-1	Sequence 1, Appl1
4	66.4	5.1	6027	4 US-09-554-467A-1	Sequence 1, Appl1
5	58.6	4.5	2523	2 US-08-410-784A-3	Sequence 3, Appl1
6	57	4.3	19124	2 US-08-487-826B-13	Sequence 13, Appl1
7	55.8	4.3	279	4 US-09-313-294A-5397	Sequence 5397, Ap
8	55.6	4.2	612	4 US-09-902-540-1357	Sequence 1357, Ap
9	55.6	4.2	1394	3 US-08-880-499-1	Sequence 1, Appl1
10	55.6	4.2	1394	3 US-08-880-499-2	Sequence 2, Appl1
11	55.4	4.2	1055	4 US-09-806-708B-23	Sequence 23, Appl1
12	55.2	4.2	2614	4 US-09-004-056-1	Sequence 1, Appl1
13	55	4.2	396	4 US-09-640-173-53	Sequence 53, Appl1
14	55	4.2	396	4 US-09-713-550-53	Sequence 53, Appl1
15	55	4.2	396	4 US-09-825-294-53	Sequence 53, Appl1
16	55	4.2	396	4 US-09-970-966-53	Sequence 53, Appl1
17	54.6	4.2	55886	4 US-09-949-016-15129	Sequence 15129, A
18	53.8	4.1	307	4 US-09-313-294A-4743	Sequence 4743, Ap
19	53.8	4.1	6027	2 US-08-968-542C-1	Sequence 1, Appl1
20	53.8	4.1	6027	4 US-09-554-467A-1	Sequence 1, Appl1
21	52.2	4.0	1141	4 US-09-806-708B-22	Sequence 22, Appl1
22	51.6	3.9	134987	4 US-09-949-016-15348	Sequence 15348, A
23	51.6	3.9	134987	4 US-09-949-016-15349	Sequence 15349, A
24	51.6	3.9	134987	4 US-09-949-016-15350	Sequence 15350, A
25	51.6	3.9	134987	4 US-09-949-016-15507	Sequence 15507, A
26	51.6	3.9	134987	4 US-09-949-016-15508	Sequence 15508, A
27	51.6	3.9	134987	4 US-09-949-016-15509	Sequence 15509, A

C	28	51.4	3.9	50383	4 US-09-949-016-17600	Sequence 17600, A
C	29	51.4	3.9	129415	4 US-09-949-016-16997	Sequence 16997, A
C	30	51	3.9	601	4 US-09-949-016-156535	Sequence 156535, A
C	31	51	3.9	1141	4 US-09-806-708B-22	Sequence 22, Appl1
C	32	50.6	3.9	1039	4 US-09-902-540-1280	Sequence 1280, Ap
C	33	50.6	3.9	16573	4 US-09-949-016-14876	Sequence 14876, A
C	34	50.6	3.9	18773	4 US-09-949-016-14164	Sequence 14164, A
C	35	50.4	3.8	731	1 US-08-451-405A-2	Sequence 2, Appl1
C	36	50.2	3.8	832	4 US-09-621-976-2813	Sequence 2813, Ap
C	37	50.2	3.8	317366	4 US-09-949-016-16001	Sequence 16001, A
C	38	50	3.8	1392	3 US-09-257-584-1	Sequence 1, Appl1
C	39	49.8	3.8	2435	4 US-09-306-593-1	Sequence 1, Appl1
C	40	49.8	3.8	231129	4 US-09-949-016-16110	Sequence 16110, A
C	41	49.8	3.8	266283	4 US-09-949-016-11934	Sequence 11934, A
C	42	49.6	3.8	640681	4 US-09-790-988-1	Sequence 1, Appl1
C	43	49.2	3.8	126176	4 US-09-949-016-16137	Sequence 16137, A
C	44	49.2	3.8	126176	4 US-09-949-016-16138	Sequence 16138, A
C	45	47.8	3.6	601	4 US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Huffman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: AND METHOD OF USING SAME

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 1311; DB 3; Length 1394;

Best Local Similarity 100.0%; Pred. No. 1, 1e-313;

Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAGGCTC 60
DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAGGCTC 60
QY 61 CTTCTTCCCTTATTACTGA CTGAATCGGGGTTTACAAAAAATTCCAGGGTGATAT 120
DB 61 CTTCTTCCCTTATTACTGA CTGAATCGGGGTTTACAAAAAATTCCAGGGTGATAT 120
QY 121 CTCATGCTGCTCTTCCGACCTCGGTGACATTTCTTGATGTCGGTGGTCCCAT 180
DB 121 CTCATGCTGCTCTTCCGACCTCGGTGACATTTCTTGATGTCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACAGACCTTTTCGGGACACCATCAAGGGCCTTTGGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACAGACCTTTTCGGGACACCATCAAGGGCCTTTGGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
DB 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
QY 301 TTATTTATCTTTGATATTTATTTTGAATAAATAAATAAATAAATAAATAAATAAATAA 360
DB 301 TTATTTATCTTTGATATTTATTTTGAATAAATAAATAAATAAATAAATAAATAAATAA 360
QY 361 GGGCCTCAGCATGATTTTCCTTAGGGCCAGAAATGCGAGGACAGCCAGCTGTATG 420
DB 361 GGGCCTCAGCATGATTTTCCTTAGGGCCAGAAATGCGAGGACAGCCAGCTGTATG 420
QY 421 TTCACCTTTGGCACTACCCGAAACAGATTTTAAAAAATAACCAAGTAACTATCCACT 480
DB 421 TTCACCTTTGGCACTACCCGAAACAGATTTTAAAAAATAACCAAGTAACTATCCACT 480
QY 481 CGAAGGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCCTTTAAAAA 540
DB 481 CGAAGGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCCTTTAAAAA 540
QY 541 CAAGCATATTTGAAAAAGACAAATTAATGTTACAGTTTACAAACATCTMAAGCGACAA 600
DB 541 CAAGCATATTTGAAAAAGACAAATTAATGTTACAGTTTACAAACATCTMAAGCGACAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATTTTGT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATTTTGT 660
QY 661 ATTGTTTATATACATTTTCTCTCTTACATAGATGATTTTCTTCGATTTTAAAA 720
DB 661 ATTGTTTATATACATTTTCTCTCTTACATAGATGATTTTCTTCGATTTTAAAA 720
QY 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGCTAGATTTCTCGTTCAAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGCTAGATTTCTCGTTCAAAAATC 780
QY 781 TTTCTGATTTTTTAAAGAGTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGAGTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTTCTCTTAAATAATGAAAAACATTTAGAAAAATGAGT 900
DB 841 TTCAAAAAAATTAAGTTATTTTCTCTTAAATAATGAAAAACATTTAGAAAAATGAGT 900
QY 901 TGCACAGTAGCCCTAGAAATGTTTTCCCAATTAATTAATCAATCACTGATTAATTTTG 960
DB 901 TGCACAGTAGCCCTAGAAATGTTTTCCCAATTAATTAATTAATCAATCACTGATTAATTTTG 960
QY 961 GCCAGGCCCTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGGCCCTAATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAAAGCTATGCTCC 1080
DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAAAGCTATGCTCC 1080
QY 1081 TGAGATGTGGGTTTGCAACGATAGCACCGTATATCATAGTGTGCTTACGTCA 1140
```

```
DB 1081 TGAGATGTGGGTTTGCAACGATAGCACCGTATATCATAGTGTGCTTACGTCA 1140
QY 1141 GGTTCGGACGCTCTGCTGATCTCATGACGATACATGCTTTTCAACCGTTGCTC 1200
DB 1141 GGTTCGGACGCTCTGCTGATCTCATGACGATACATGCTTTTCAACCGTTGCTC 1200
QY 1201 TTGTTTCATGCTGCAAGCCTTGCTTATCTGTAACCAAGAGGATACCTACTCCCAACAT 1260
DB 1201 TTGTTTCATGCTGCAAGCCTTGCTTATCTGTAACCAAGAGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTCATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garneat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREMITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-2

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No 1.1e-313; Indels 0; Gaps 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAGGCTC 60
DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAGGCTC 60
QY 61 CTTCTTCCCTTATTACTGA CTGAATCGGGGTTTACAAAAAATTCCAGGGTGATAT 120
DB 61 CTTCTTCCCTTATTACTGA CTGAATCGGGGTTTACAAAAAATTCCAGGGTGATAT 120
```


QY 121 CTCATGTTCCACTCTTCCACCTCGGTTGCCAATTTCTGATGTCGGTGTTCAT 180
DB 121 CTCATGTTCCACTCTTCCACCTCGGTTGCCAATTTCTGATGTCGGTGTTCAT 180
QY 181 CTGACCCGAGGCCCTTCAGACACCTTTGGGCAACCCATCAAGGGCTTTGSAATGCCCA 240
DB 181 CTGACCCGAGGCCCTTCAGACACCTTTGGGCAACCCATCAAGGGCTTTGSAATGCCCA 240
QY 241 CGAGAGCTATCGGGCTGTGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
DB 241 CGAGAGCTATCGGGCTGTGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
QY 301 TTATTTATCTTTAGATATTTATTTAATTTTGGAAAAATTAACAACTTATCTTTGGTGA 360
DB 301 TTATTTATCTTTAGATATTTATTTAATTTTGGAAAAATTAACAACTTATCTTTGGTGA 360
QY 361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGGACGACCATGTCTAGTG 420
DB 361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGGACGACCATGTCTAGTG 420
QY 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAACTAATCCACT 480
DB 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAACTAATCCACT 480
QY 481 CGAAAGCTATCATGTATATTTTAAAGAAACATCTATTTAAACCAAGTAACTAATCCACT 540
DB 481 CGAAAGCTATCATGTATATTTTAAAGAAACATCTATTTAAACCAAGTAACTAATCCACT 540
QY 541 CAAGCATATTTTCGAAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600
DB 541 CAAGCATATTTTCGAAAGAGACAAATTTATTTACAGTTTACAAACATCTAAGAGGACAA 600
QY 601 TTATATCGAAAGGTAGACCTATGACGTTCAAGTTTTCTTTTCAATCTTGTATTTTGT 660
DB 601 TTATATCGAAAGGTAGACCTATGACGTTCAAGTTTTCTTTTCAATCTTGTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTCTCTTAAATAGATGATTTTCTTCCGATTTTATATA 720
DB 661 ATTGTTTTTATATACATTTTCTCTCTTAAATAGATGATTTTCTTCCGATTTTATATA 720
QY 721 ATGACTATTAAGTCATTTTATATATAGAGACGATCTCGTAGATTTCTGTTCAAAAATC 780
DB 721 ATGACTATTAAGTCATTTTATATAGAGACGATCTCGTAGATTTCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTGTTCTTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTGTTCTTTCAAGAAATTTGATTTT 840
QY 841 TTTCAAAAAAATTAAGTTATTTTCTCTTTAATAATGAAAAACATTAAGAAAAATAGAGT 900
DB 841 TTTCAAAAAAATTAAGTTATTTTCTCTTTAATAATGAAAAACATTAAGAAAAATAGAGT 900
QY 901 TGCAGAGCTAGCCCTAGATGTTTTTCCCAATTAATCACTGCTGTAATTAATTTT 960
DB 901 TGCAGAGCTAGCCCTAGATGTTTTTCCCAATTAATCACTGCTGTAATTAATTTT 960
QY 961 GCCAGGCCCATTAATTTTAAACCGAACTGAAATGAGGAAACCAAACTGAGCTAT 1020
DB 961 GCCAGGCCCATTAATTTTAAACCGAACTGAAATGAGGAAACCAAACTGAGCTAT 1020
QY 1021 TTTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
DB 1021 TTTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAACTATTTGCC 1080
QY 1081 TGAGATGTGCGGTTTGGCAAGATGCGATCATAGTCTATGAGTGTGCTTCAAGCTCA 1140
DB 1081 TGAGATGTGCGGTTTGGCAAGATGCGATCATAGTCTATGAGTGTGCTTCAAGCTCA 1140
QY 1141 GGTTCGGACGCTCTGTGTCACTGACATGCGCATATCACTGCTTGTCAACCGTTGCTC 1200
DB 1141 GGTTCGGACGCTCTGTGTCACTGACATGCGCATATCACTGCTTGTCAACCGTTGCTC 1200

QY 1201 TTGTCATGTCGCAAGCCCTTGCTATTCGAAACCAAGGATACCTACTCCCAACAT 1260
DB 1201 TTGTCATGTCGCAAGCCCTTGCTATTCGAAACCAAGGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 3

US-08-968-542C-1/C
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dutil Codes For A No. 5981728el Search
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/968,542C
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm
; IMMEDIATE SOURCE:
; LIBRARY: (GFI1
; CLONE: pMgfi10; pMg6Aa; pMg6-2M
; US-08-968-542C-1

Query Match 5.1%; Score 66.4; DB 2; Length 6027;

Best Local Similarity 77.5%; Pred. No. 2.4e-06;

Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTTGGCAACCTGTGTTCTTTCAAGAAATTTGATTTTTCAAAAAAA 851
DB 5606 TCTACGGGTAGTTTGGAAACCCCAATTT-TTCCAAGGATTTTCATTTTTCAGAGAAAA 5548
QY 852 TTAGTTATTTTCTCTTTTAAATAAGAAAAACCTTAGAAAAATAGAGTTGCCAGACTAG 911

Db 5547 TTAGTTATTTTTCATTGGAAAAATGAAATCTTTGGAAAAATAGATTACATACTAG 5488

RESULT 4
US-09-554-467A-1/C
Sequence 1, Application US/09554467A
Patent No. 6639125
GENERAL INFORMATION:
APPLICANT: Myers, Alan M.
APPLICANT: James, Martha G.
TITLE OF INVENTION: dulla Coding for a No. 6639125el Starch Synthase and Uses
FILE REFERENCE: D6036PCT
CURRENT APPLICATION NUMBER: US/09/554,467A
CURRENT FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: PCT/US98/24225
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 08/062,102
PRIOR FILING DATE: 1997-11-12
NUMBER OF SEQ ID NOS: 37
SEQ ID NO 1
LENGTH: 6027
TYPE: DNA
ORGANISM: maize
FEATURE:
OTHER INFORMATION: cdna sequence corresponding to the gene encoding the
OTHER INFORMATION: starch synthase enzyme DUL.
US-09-554-467A-1

Query Match
Best Local Similarity 5.1%; Score 66.4; DB 4; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2,4e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
Db 5606 TCTACGGGCTAGTATGCGAACCCCATTT-TTCCAGAGGATTTCCATTTTCCAGAAAA 5548

QY 852 TTAGTTATTTTCTCTTTATAAATAGAAAAACATTAGAAAAATAGGTCAGACTAG 911
Db 5547 TTAGTTATTTTCTCTTTATTTGAAAAATGAAATCTTTGGAAAAATAGATTACATACTAG 5488

RESULT 5
US-08-410-784A-3/C
Sequence 3, Application US/08410784A
Patent No. 5912413
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Heine, Ph.D., Holliday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-410-784A-3

Query Match
Best Local Similarity 4.5%; Score 58.6; DB 2; Length 2523;
Best Local Similarity 71.4%; Pred. No. 0.00015;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 774 AAAATCTTCTGATTTTAAAGAGTAGTTGGCAACCTGTTCTTCAAGAAATT 833
Db 279 ACATATACATATACATATATAGCTTAGTTGACATCTCATTTTACAAAG--TTT 222

QY 834 TGATTTTCAAAAAAATTAAGTTATTTTCTCTTTATAAATAGAAAAACATTAGAAAA 893
Db 221 TACATTTTCAAAATTAATAGTTATTTTCTCTTGA-AAATAGAAATTTCTCAGAAAA 163

QY 894 ATAGAGTCCGAGTACGCCCTAGAT 920
Db 162 ATAGAGTTACAACTAGCTTAAAT 136

RESULT 6
US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim J.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

Query Match 4.3%; Score 57; DB 2; Length 19124;
 Best Local Similarity 50.2%; Pred. No. 0.00078;
 Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAAAGTAAAGACGTTGACATTTTCTTTTCAATCTGTATTTGTTATTTTATAT 673
 DB 15677 TAAAGTAAAGACGTTGACATTTTCTTTTCAATCTGTATTTTATAT 15618
 QY 674 ACATTTCTCTCTTACAAATAGAGTATTTCTCCATTTTATAAAGTAAAGT 733
 DB 15617 AAATTTTCTCTTACAAATAGAGTATTTCTCCATTTTATAAAGTAAAGT 15558
 QY 734 CATTTTATATAGAGACGACGATCTGATTTCTGCAAAAATCTTCTGATTTT 793
 DB 15557 TATTTTATATATAGAGACGACGATCTGATTTCTGCAAAAATCTTCTGATTTT 15498
 QY 794 TAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTTCAAAAAAAT 853
 DB 15497 TTTTATTTTAAATATTTTCTTTTATATTTTCAATTTCTTTTCAATTTTAAATTT 15438
 QY 854 AGTTTATTTCTCTTTTAAATAGAAAACACTTGAAGAAA 894
 DB 15437 GTTTTATATTTCTTTTAAATAGAAAACACTTGAATATAA 15397

RESULT 7
 US-09-313-294A-5397
 ; Sequence 5397, Application US/09313294A
 ; Patent No. 6476212
 ; GENERAL INFORMATION:
 ; APPLICANT: Lalguadi, Raghunath V.
 ; APPLICANT: Ito, Laura Y.
 ; APPLICANT: Sherman, Bradley K.
 ; FILE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
 ; TITLE REFERENCE: PL-0017 US
 ; CURRENT APPLICATION NUMBER: US/09/313,294A
 ; CURRENT FILING DATE: 1999-05-14
 ; NUMBER OF SEQ ID NOS: 7600
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 5397
 ; LENGTH: 279
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: incycle ID No. 6476212 700350078H1
 ; NAME/KEY: unsure
 ; LOCATION: 10, 12, 185-186, 204, 253, 274, 278
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-313-294A-5397

Query Match 4.3%; Score 55.8; DB 4; Length 279;
 Best Local Similarity 69.0%; Pred. No. 0.00031;
 Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
 QY 789 TTTTAAAGACGTTGACATTTTCTTTTCAAGAAATTTGATTTTCAAAA 848
 DB 3 TTCTTAANGCCTAGTTGGAAACCCATTTTCCCGCGGTTTCAATTTTCCCAAGG 62
 QY 849 AAATAGTATTTTCTCTTAT-AAATAGAAAACCTAGAAAATAGAGTTGCAGA 907

DB 63 AAGTAGAACATTTTCCCTTGGGAAATAGAAATCTTTGGGAAATCGAGTTCCCAA 122
 QY 908 CTAGCCCTA 916
 DB 123 CTAGCCCTA 131

RESULT 8
 US-09-902-540-1357/c
 ; Sequence 1357, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; FILE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; TITLE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 1357
 ; LENGTH: 612
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)-(612)
 ; OTHER INFORMATION: unsure at all n locations
 US-09-902-540-1357

Query Match 4.2%; Score 55.6; DB 4; Length 612;
 Best Local Similarity 51.4%; Pred. No. 0.00047;
 Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCAATCTGTTATTTGTTATTTGTTTATATACATTTCTCT 685
 DB 560 TTTTATATTTTCTTTTCAATCTGTTATTTGTTATTTGTTTATATACATTTCTCT 501
 QY 686 CTGACATAGAGATTTCTCCGATTTTATTAATAAGCTAATGATTTTATATA 745
 DB 500 TTTTATATTTTCTTTTCAATCTGTTATTTGTTATTTGTTTATATATTTTATA 441
 QY 746 AGAGCAGCATGTGAGATTTCTGTTCAAAAATCTTCTGATTTTAAAGCTAGT 805
 DB 440 ATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 381
 QY 806 TGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTATTTCT 865
 DB 380 TTTTATATTTTCTTTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTATTTCT 321
 QY 866 CTTTATA 872
 DB 320 TATTTTA 314

RESULT 9
 US-08-880-499-1/c
 ; Sequence 1, Application US/08880499
 ; Patent No. 6037523
 ; GENERAL INFORMATION:
 ; APPLICANT: Albertson, Marc C.
 ; APPLICANT: Fox, Tim W.
 ; APPLICANT: Carl, Garnaat W.
 ; APPLICANT: Huffman, Gary A.
 ; APPLICANT: Kendall, Timmy U.
 ; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
 ; TITLE OF INVENTION: AND METHOD OF USING SAME
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 4.2%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;

Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGAGTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTTGGAACATCTAGGCTAGCTGGCAACTATTTT 891
QY 821 TTTCAGAGATTTGATTTTTCAGAAAATTAAGTTATTTCTCTTATAAATGAA 880
DB 890 TCTAAGTGTCTTATTTTAAAGAGAAATTAACCTAATTTTTTGAATAACAAA 831
QY 881 AACCTTAGAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTTCGCAACTAGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACTGTGTA 950
DB 770 CGAGAACTCA 761

RESULT 10

US-08-880-499-2/c
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garraat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Jimmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 4.2%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;

Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGAGTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTTGGAACATCTAGGCTAGCTGGCAACTATTTT 891
QY 821 TTTCAGAGATTTGATTTTTCAGAAAATTAAGTTATTTCTCTTATAAATGAA 880
DB 890 TCTAAGTGTCTTATTTTAAAGAGAAATTAACCTAATTTTTTGAATAACAAA 831
QY 881 AACCTTAGAAAATAGAGTTGCCAGACTAGCCCTAGATGTTTCCCATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTTCGCAACTAGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACTGTGTA 950
DB 770 CGAGAACTCA 761

RESULT 11

US-09-806-708B-23
Sequence 23, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1998-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin version 3.0
SEQ ID NO 23
LENGTH: 1055
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)..(1055)
OTHER INFORMATION: consensus sequence of A.T. and L.A. FAEI promoters
US-09-806-708B-23

Query Match 4.2%; Score 55.4; DB 4; Length 1055;
Best Local Similarity 22.5%; Pred. No. 0.00065;
Matches 182; Conservative 181; Mismatches 430; Indels 15; Gaps 3;


```

QY      292 TCACATATATATATCTTACATATATATATATTTTGGAAAAATAACAACTTATAC 351
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      37 YCANNTGKRCYARWGMWTTAYWTTATKWTGTMAMTWAMKTRKMCSTAMNNANW 96
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      352 TTTTGTGTAGGGCCCTCAG-CATAGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGC 410
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      97 TTTCATARKMTGTGRMTKTNNAATGTRMTGTYMTNNNGSTYTMARRYTRTRWCTYAM 156
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      411 ATGTCTAGTGTCACTATTTGGCACTACCCAGAAACAAGATTAAAAATAACCAAAATGA 470
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      157 MYGASVAGNASTRTTYTRWKMCKKKSABATRGARVYAMVYAMARTRGWTAKMA 216
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      471 CTATCCACTCGAAAGCTATCATGTATGT---TTAAGAAACCTCTATTAAAAACGCA 527
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      217 YAAWYMNNNNNNAKACKRAITWGRKSNCTCTTAGTGTTCATCMAYTCCAGATATKK 276
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      528 TCCTCTTAAAAAACAAGCATATTTTGGAAAGAGACAATTTGTACAGTTTACAAACATC 587
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      277 WKTKWTSAAAGMTWNNNNNNNTTTTKAATYAAARMMWMSATTTWMAAMTSKRTWYGR 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      588 TAAAGCGACAAATTAATGCAAGAGTAAGCTATGACGTTACAGATTTTCTTTTCATTC 647
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      337 KTMANNNGTTCWTRWMAWTKMMKTKGTTNNNGGRTYTGWTTKXATTTTAANNC 396
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      648 TTGTTATTTTGTATTGTTTATATACATTTCTCTTCAATAGAGTGAATTTCTT 707
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      397 TTAAWKKCTCTMNNNTTAAKATTYWATCYWKSMTNGTSRYAARYYTWYAMTRRYANN 456
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      708 CGATTTTAAATAAGCTATA-----AAGCATTTTATATAAAGACAGCAT 756
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      457 NNTKTWKACTWTTTKCTTANNNTAAWTKSSANCTSRTRKTKNCWRAGSKTASMGRA 516
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      757 GTCTGATATCTGCTCAAAAATCTTCTGATTTTAAAGCTAGTTTGGCAACCTG 816
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      517 YAAWYTGWMTNAMAAMWCTWYVYRAGAAMTAMWMTSATCVCATATATTAGCAAGGS 576
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      817 TTTCTTTCAAGAAATTTGATTTTTCAAAAAAATTAAGTTTATTTCTTTATAAAT 876
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      577 TAGNNNNNNNNNCAATCAATCAATCAASACAMANNATTCCTAANNATANNNTGCV 636
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      877 AGAAACACTTAGAAAAATAGAGTGCACACTAGCCCTGAAGTGTTCCTCAATTAAT 936
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      637 NATKTATMTNNNNNNNAGTWTNNNNNNNMAASATYTAATAATATKATNTTAMAG 686
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      937 ACAATCACTGTATATATTTTGGCCAGCCCATTAATTTTAAACCGAACTGAAT 996
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      697 AYABAAAYTTRTANNGACTTTTNTTGTGMRNTTAAABGMANNNNNNNNNNNGACMA 756
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      997 CGAGCGAAACCAATCTGAGCTATTTCTCTAGATTAAGTAAAAAGAGAGAGAGAG 1056
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      757 WRTTATANGTNNNNNNNNNNNNAATTTTATTTTWTTRKANNNNNNNNNAAYYGAAM 816
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      1057 AATCAGTTTAAAGTATGTCCTGAG 1084
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      817 KNNITCMTCKAMKAAATGAATTTNAG 844
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
US-09-004-056-1/C
; Sequence 1, Application US/09004056A
; Patent No. 656586
; GENERAL INFORMATION:
; APPLICANT: Calsgene LLC
; TITLE OF INVENTION: Plant Expansion Promoter Sequences
; FILE REFERENCE: 125
; CURRENT APPLICATION NUMBER: US/09/004,056A
; EARLIER FILING DATE: 1998-01-07
; EARLIER APPLICATION NUMBER: 60034914
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```

```

; LENGTH: 2614
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (930)
; OTHER INFORMATION: unknown nucleotide
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (947)
; OTHER INFORMATION: unknown nucleotide
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (956)
; OTHER INFORMATION: unknown nucleotide
US-09-004-056-1

Query Match      4.2%; Score 55.2; DB 4; Length 2614;
Best Local Similarity 54.2%; Pred. No. 0.001; Indels 3; Gaps 2;
Matches 155; Conservative 0; Mismatches 128; Indels 3; Gaps 2;

QY      611 AGGTAGCTATGACGTTGAGATTTTCTTTTCATCTGTATTTTGTATTTGTTTAA 670
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      447 AGTTTAGTTGGTTATCAATTTTTCATATTAATTTAGTTTATTTCTAATTTA 388
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      671 TATACATTTCTCTCTTAACAATGAGTGATTTCTCCGATTTTATAAATGACTATTA 730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      387 TGTTGCAAAAGAACTTTATTTATATTTTAAATATATATGATTAAT-TTAAAG 329
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      731 AGTCATTTTATATAGACAGCATGCTGATCTGTCGAAATCTTCTGATTT 790
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      328 TATTTTCAATATATATATTTCAAGAAACAATATTTTGAATTTTGTGATTT 269
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      791 TTTAAGACTAGTTTGGCAACCTGTTCTTCAAGATTTTGAATTTTCAAAAAA 850
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      268 TTTAAATGTAGTATTTTATTTCTTATTTATATGATATAT-TTATTTAGAAAAA 211
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      851 ATTAGTTATTTCTCTTATATAAATAGAAAACTTAGAAAAATA 896
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      210 ATTAGGTTAATGAAATTTAAATTTTAAATTAATTAATAAAAAATA 165
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-640-173-53
; Sequence 53, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; EARLIER FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match      4.2%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00056;
Matches 159; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY      632 TTTTCTTTTCATCTTGTATTTGTATGTTTATATACATTTTCTCTCTTACA 691
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      11 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 70
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
QY 692 ATAGAGTATTTCTCCGATTTTATATAAGTAAAGTATTTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 130
QY 752 CGCATGCTGATGATCTCGTCAAAAATCTTTGATTTTATAGAGTATTTGGCAA 811
DB 131 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 190
QY 812 CCTGTTCTTTCAAAAGATTTGATTTTTCAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 250
QY 872 AAAATGAAAAACACTTAGAAAAATAGAGTTCAGACCTAGATGTTTCCCAAT 931
DB 251 AATTCAAAAAAGAAAAAGAAAATTAANANNANNNNNNNNNNNNNNNNNNNNN 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNNTNNNNNNANNGG 327

RESULT 14
US-09-713-550-53
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match 4.2%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00056;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCTTTTCTTTCTGTTATTTGTTATGTTTATATACATTTCTCTCTACA 691
DB 11 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 70
QY 692 ATAGAGTATTTCTCCGATTTTATATAAGTAAAGTATTTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 130
QY 752 CGCATGCTGATGATCTCGTCAAAAATCTTTGATTTTATAGAGTATTTGGCAA 811
DB 131 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 190
QY 812 CCTGTTCTTTCAAAAGATTTGATTTTTCAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 250
QY 872 AAAATGAAAAACACTTAGAAAAATAGAGTTCAGACCTAGATGTTTCCCAAT 931
DB 251 AATTCAAAAAAGAAAAAGAAAATTAANANNANNNNNNNNNNNNNNNNNNNNN 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNNTNNNNNNANNGG 327
```

```
RESULT 15
US-09-825-294-53
; Sequence 53, Application US/09825294
; Patent No. 6710170
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Stolk, John A.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steven P.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C5
; CURRENT APPLICATION NUMBER: US/09/825,294
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53
```

```
Query Match 4.2%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00056;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCTTTTCTTTCTGTTATTTGTTATGTTTATATACATTTCTCTCTACA 691
DB 11 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 70
QY 692 ATAGAGTATTTCTCCGATTTTATATAAGTAAAGTATTTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 130
QY 752 CGCATGCTGATGATCTCGTCAAAAATCTTTGATTTTATAGAGTATTTGGCAA 811
DB 131 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 190
QY 812 CCTGTTCTTTCAAAAGATTTGATTTTTCAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 250
QY 872 AAAATGAAAAACACTTAGAAAAATAGAGTTCAGACCTAGATGTTTCCCAAT 931
DB 251 AATTCAAAAAAGAAAAAGAAAATTAANANNANNNNNNNNNNNNNNNNNNNNN 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNNTNNNNNNANNGG 327
```

Search completed: September 15, 2005, 08:24:56
UDD time : 200.304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11 ; Search time 853.804 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Perfect score: 1311
Sequence: 1 cccatgctgctctcatgaaaa.....cgcacatagcttcctcgaaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 3331285599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	20	US-10-713-381-1
2	1311	100.0	1394	20	US-10-713-381-2
3	160.6	12.3	255	20	US-10-713-381-9
4	146	11.1	158	20	US-10-713-381-3
5	80.8	6.2	1261	20	US-10-425-115-14230
6	74	5.6	320	18	US-10-425-114-23340
7	74	5.6	624	18	US-10-425-114-16264

8	73	5.6	1326	20	US-10-425-115-141826	Sequence 141826, A
9	72.8	5.6	2445	18	US-10-425-114-32493	Sequence 32493, A
10	72.8	5.6	2729	20	US-10-425-115-83293	Sequence 83293, A
11	72.4	5.5	1203	20	US-10-425-115-51470	Sequence 51470, A
12	71.8	5.5	928	18	US-10-425-114-17816	Sequence 17816, A
13	71.8	5.5	1123	20	US-10-425-115-142853	Sequence 142853, A
14	71.8	5.5	2537	18	US-10-425-114-31957	Sequence 31957, A
15	71.8	5.5	2537	20	US-10-425-115-52216	Sequence 52216, A
16	71.8	5.5	3607	18	US-10-425-114-31061	Sequence 31061, A
17	71.8	5.5	3691	20	US-10-425-115-52219	Sequence 52219, A
18	70.2	5.4	960	14	US-10-198-846-6381	Sequence 6381, Ap
19	69.8	5.3	13400	21	US-10-343-477A-63	Sequence 63, Appl
20	69.6	5.3	724	20	US-10-425-115-87756	Sequence 87756, A
21	69.2	5.3	527	18	US-10-425-114-4041	Sequence 4041, Ap
22	68.8	5.2	610	20	US-10-425-115-47165	Sequence 47165, A
23	68.4	5.2	1215	18	US-10-425-114-24656	Sequence 24656, A
24	68.4	5.2	2863	20	US-10-425-115-75310	Sequence 75310, A
25	67.8	5.2	512	20	US-10-425-115-49781	Sequence 49781, A
26	67.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
27	67.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
28	67.4	5.1	1260	18	US-10-425-114-30881	Sequence 30881, A
29	67.4	5.1	1411	20	US-10-425-115-149304	Sequence 149304, A
30	67.2	5.1	2232	18	US-10-425-114-16607	Sequence 16607, A
31	67.2	5.1	2249	18	US-10-425-114-20264	Sequence 20264, A
32	66.4	5.1	6027	18	US-10-634-262-1	Sequence 1, Appl
33	66.4	5.1	6027	19	US-10-109-048-1145	Sequence 1145, Ap
34	66.4	5.1	6051	20	US-10-425-115-41589	Sequence 41589, A
35	66.2	5.0	439	9	US-09-960-352-1009	Sequence 1009, Ap
36	66	5.0	1376	18	US-10-425-114-2073	Sequence 2073, Ap
37	65.6	5.0	721	20	US-10-425-115-177935	Sequence 177935, A
38	65.6	5.0	2274	18	US-10-425-114-777	Sequence 777, App
39	65.6	5.0	2274	18	US-10-425-114-830	Sequence 830, App
40	65.6	5.0	2632	20	US-10-425-115-162957	Sequence 162957, A
41	65.6	5.0	8056	20	US-10-473-126-386	Sequence 386, App
42	65	5.0	1364	20	US-10-425-115-150820	Sequence 150820, A
43	64.8	4.9	833	20	US-10-425-115-30824	Sequence 30824, A
44	64.8	4.9	1838	20	US-10-425-115-21012	Sequence 21012, A
45	64.2	4.9	581	20	US-10-425-115-46324	Sequence 46324, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OR INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTCATGAAAAGATGATGATGCTCTATATCCGTTTCTTAGGCTCC 60


```
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCATATCCGTTTCTTAAGGTCC 60
Qy 61 CTTCTTCTGCTTTATTACTGACTGAATCGGGGTTTACAAAACTTCCACGGGTGAT 120
Db 61 CTTCTTCTGCTTTATTACTGACTGAATCGGGGTTTACAAAACTTCCACGGGTGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTGGGTGGCACTTCTTGGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTGGGTGGCACTTCTTGGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGGCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
Qy 301 TTATATCTTTAGATATTAATTTTGGAAAAATACAACTTATCTTTGTTGTA 360
Db 301 TTATATCTTTAGATATTAATTTTGGAAAAATACAACTTATCTTTGTTGTA 360
Qy 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCATG 420
Db 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCATG 420
Qy 421 TCACATTTGGCTACCCAGAACAAAGTTTAAAAAATACCAAGTAACTATCCACT 480
Db 421 TCACATTTGGCTACCCAGAACAAAGTTTAAAAAATACCAAGTAACTATCCACT 480
Qy 481 CGAAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTTAAAAA 540
Db 481 CGAAAGCTATCATGATGTTTAAAGAAACATCTATTTAAACACGATCCTTAAAAA 540
Qy 541 CAAGCATATTTGAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
Qy 601 TTATATCGAAAGGTAGCTATGACGTTCAATTTTCTTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAAGGTAGCTATGACGTTCAATTTTCTTTTCAATCTGTTATTTGTT 660
Qy 661 ATGTTTTTATATACATTTCTCTCTTACATAGAGTATTTCTTCGATTTTAA 720
Db 661 ATGTTTTTATATACATTTCTCTCTTACATAGAGTATTTCTTCGATTTTAA 720
Qy 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGTCGTAGATTCTCGTTCAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGTCGTAGATTCTCGTTCAAAATC 780
Qy 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTATATTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGTATATTTG 960
Qy 961 GCGAGCCCAATAAATTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACTAT 1020
Db 961 GCGAGCCCAATAAATTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCTTCCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCTTCCC 1080
Qy 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTAATCATAGTCTATAGTGCCTACGTCA 1140
Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTAATCATAGTCTATAGTGCCTACGTCA 1140
```

```
Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTAATCATAGTCTATAGTGCCTACGTCA 1140
Qy 1141 GGTGGGAGCTCTGTGTCATCTCAATGGGATCTACATGCTTTGTTCAACCGTTCGTC 1200
Db 1141 GGTGGGAGCTCTGTGTCATCTCAATGGGATCTACATGCTTTGTTCAACCGTTCGTC 1200
Qy 1201 TTGTTCAATCGTCGAAGCCCTTCTTCTGAAACCAAGAGATGATCTACTCCCAACAT 1260
Db 1201 TTGTTCAATCGTCGAAGCCCTTCTTCTGAAACCAAGAGATGATCTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACGACATATGTTTCTGAAAC 1311
Db 1261 CCATCTTACTCATGCACTTCCATGCAACGACATATGTTTCTGAAAC 1311

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCATATCCGTTTCTTAAGGTCC 60
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTCATATCCGTTTCTTAAGGTCC 60
Qy 61 CTTCTTCTGCTTTATTACTGACTGAATCGGGGTTTACAAAACTTCCACGGGTGAT 120
Db 61 CTTCTTCTGCTTTATTACTGACTGAATCGGGGTTTACAAAACTTCCACGGGTGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTGGGTGGCACTTCTTGGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTGGGTGGCACTTCTTGGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGGCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGTCACCTATA 300
Qy 301 TTATATCTTTAGATATTAATTTTGGAAAAATACAACTTATCTTTGTTGTA 360
Db 301 TTATATCTTTAGATATTAATTTTGGAAAAATACAACTTATCTTTGTTGTA 360
Qy 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCATG 420
Db 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCATG 420
Qy 421 TCACATTTGGCTACCCAGAACAAAGTTTAAAAAATACCAAGTAACTATCCACT 480
Db 421 TCACATTTGGCTACCCAGAACAAAGTTTAAAAAATACCAAGTAACTATCCACT 480
```



```

Db 421 TCCACTTTGGCACTACCCAGAACAGATTTAAABAAAATACCAAGTAATCAATCCACT 480
Qy 481 CGAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAAATGTTTAAAGAAACATCTATTAAACCAAGTCTCTTAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATTTATGTACGTTTAAACATCTAAGACGCAAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATTTATGTACGTTTAAACATCTAAGACGCAAA 600
Qy 601 TTATATCGAAGAGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGT 660
Db 601 TTATATCGAAGAGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTTAAATAGATGATTTTCTTCCGATTTTATAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTTAAATAGATGATTTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATAAGTCAATTTTATATTAAGAGACGATGCTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAGTCAATTTTATATTAAGAGACGATGCTAGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTATTTTCTTTTAAATAGAAACATTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTATTTTCTTTTAAATAGAAACATTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTGAATGTTTCCCAATTAATTCATCTGCTGTTAATTTATTTG 960
Db 901 TGGCAGACTAGCCCTGAATGTTTCCCAATTAATTCATCTGCTGTTAATTTATTTG 960
Qy 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGACGCAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATGACGCAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCACTTTTAAGCTATTTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCACTTTTAAGCTATTTGCC 1080
Qy 1081 TGAAGATGCGGTTTGGCAACGATAGCAGCCTATATCTATGCTCATAGGCTTACGTCA 1140
Db 1081 TGAAGATGCGGTTTGGCAACGATAGCAGCCTATATCTATGCTCATAGGCTTACGTCA 1140
Qy 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCAATGCTGTTGTTCAACCGTTGCTC 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCAATGCTGTTGTTCAACCGTTGCTC 1200
Qy 1201 TTGTTCCATGCTCAAGCCTTGGCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
Db 1201 TTGTTCCATGCTCAAGCCTTGGCTTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
Qy 1261 CCATCTTACGATGCAACTTCCATGCAAAACGCAATATGTTTCTGAAAC 1311
Db 1261 CCATCTTACGATGCAACTTCCATGCAAAACGCAATATGTTTCTGAAAC 1311

```

RESULT 3
US-10-713-381-9

```

; Sequence 9, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R

```

```

; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 12.3%; Score 160.6; DB 20; Length 255;
Best Local Similarity 97.6%; Pred. No. 6.1e-25;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1145 CGGAGCTCTCGTGCATCTCAACAGGATACATACATGCTTGTCAACGTTGCTTGT 1204
Db 5 CCGCGATCCGTCATCTCAACAGGATACATACATGCTTGTCAACGTTGCTTGT 64
Qy 1205 TCCATGCTCAAGCTTGCCTATTCTGAACCAAGAGATACCTATCCCAACATTCAT 1264
Db 65 TCCATGCTCAAGCTTGCCTATTCTGAACCAAGAGATACCTATCCCAACATTCAT 124
Qy 1265 CTACTCATGCAACTTCATGCAACAGCAGATATGTTTCTGAAAC 1311
Db 125 CTACTCATGCAACTTCATGCAACAGCAGATATGTTTCTGAAAC 171

```

RESULT 4
US-10-713-381-3

```

; Sequence 3, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

```

Query Match 11.1%; Score 146; DB 20; Length 158;

Best Local Similarity 99.4%; Pred. No. 7.3e-22;

Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

Qy 1155 CGTGCATCTCAATGAGCATATCTACATGCTTGTCAACGTTGCTC-TTGTTCATGCTC 1213
Db 1 CGTGCATCTCAATGAGCATATCTACATGCTTGTTCACCGTTGCTTGTTCATGCTC 60

```

```

Qy 1214 CAAGCTTGCTATTCTGAACCAAGAGATACCTATCCCAACATTCATCTCAT 1273
Db 61 CAAGCTTGCTATTCTGAACCAAGAGATACCTATCCCAACATTCATCTCAT 120

```

```

Qy 1274 GCACTTCATGCAACAGCAGCATATGTTTCTGAAAC 1311
Db 121 GCACTTCATGCAACAGCAGCATATGTTTCTGAAAC 158

```

RESULT 5
US-10-425-115-134230


```
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MR14577_53903C.1
US-10-425-115-134230
```

```
Query Match          6.2%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 3.1e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
    |||||
DB 999 TCTTAAGGCTAGTTGAAACCTTATTTTAAAGATTTTCTTTTAAAGTAA 1058
    |||||
QY 852 TTAGTTATTTTCTCTTATTAATAATGAAAAACATTGAAAAATAGAGTCCAGACTAG 911
    |||||
DB 1059 TTAGTTATTTTCTCTTGAATAATGAAATCTCTTGAAAAATGAGTTGCTAACTAG 1118
    |||||
QY 912 CCCT 915
    |||||
DB 1119 CCCT 1122
```

```
RESULT 6
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_F11
US-10-425-114-23340
```

```
Query Match          5.6%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 62 TTTTGGGCTAGTTGAAATCTCATTTTTCAGAGATTTTATTTTCCAAAGAA 121
    |||||
QY 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAAAATGAGTTGCCAGACTA 910
    |||||
```

```
DB 122 ATTAGTTATTTCTCTTGAATAATGAAATTCCTTGGAATTAAGATTTCTAAACGA 181
QY 911 GCCCTAGAT 920
    |||||
DB 182 GCCCTTAAT 191
```

```
RESULT 7
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 16264
/ LENGTH: 624
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3062-023-G10_F11
US-10-425-114-16264
```

```
Query Match          5.6%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 6.7e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
QY 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 374 TTTTGGGCTAGTTGAAATCTCATTTTTCAGAGATTTTATTTTCTTAAGAA 433
    |||||
QY 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAAAATAGAGTCCAGACTA 910
    |||||
DB 434 ATTAGTTATTTCTCTTGAATAATGAAATTCCTTGGAATTAAGATTTCTAAACGA 493
    |||||
QY 911 GCCCTAGAT 920
    |||||
DB 494 GCCCTTAAT 503
```

```
RESULT 8
US-10-425-115-141826
/ Sequence 141826, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 141826
/ LENGTH: 1326
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MR14577_60829C.1
US-10-425-115-141826
```

```
Query Match          5.6%; Score 73; DB 20; Length 1326;
```


Page 7

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 16:07:32 ; Search time 4088.83 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1_COPY_1_1311

Perfect score: 1311
Sequence: 1 ccatggtcgtctatgaaaa.....cgcatatgcttctcgaac 1311

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hlc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	908	69.3	963	9	CC656933 OGMDO20TM
2	679	51.8	915	9	CG224225 OGIAG08TV
3	419.8	32.0	687	9	CC656939 OGMDO20TV
4	96	7.3	715	9	CG252571 CGAB057C
5	92	7.0	967	9	CL235046 ZMMBB057
6	88	6.7	814	9	CG048704 PU1019TB
7	87.8	6.7	754	9	CG414922 ZMMBB029
8	86.8	6.6	950	8	CC439901 PUHRLV5TB
9	86.6	6.6	652	8	CC384247 PUHOC67TB
10	86.6	6.6	797	8	CC400575 PUHLU61TD
11	86.6	6.6	820	8	CC400574 PUHLU61TB
12	85.2	6.5	471	9	CG103452 PUHBL9TB
13	85.2	6.5	765	9	CG082135 PUFOX12TD
14	85.2	6.5	781	9	CC630219 OGUCC53TV
15	85.2	6.5	781	9	CC630219 OGUCC53TV
16	85.2	6.5	815	9	CC630219 OGUCC53TV
17	85	6.5	789	8	CC433618 PUHHP17TD
18	84.6	6.5	1092	8	CNS020K7
19	84	6.4	793	8	BZ786381
20	84	6.4	1078	9	CL957678
21	83.8	6.4	530	9	CG201774 PUICH24TB
22	83.6	6.4	722	9	CG333914 OGUAD14TB
23	83.6	6.4	722	9	CG333929 OGUAD14TB
24	83.6	6.4	781	9	CG034985 PUIGR68TB

25	83.6	6.4	861	8	BZ797976
26	83.6	6.4	865	8	CC430754
27	81.8	6.2	861	9	CG102092
28	81.8	6.2	947	8	CC435780
29	81.8	6.2	981	8	BZ784278
30	81.8	6.2	1017	9	CC620594
31	81.6	6.2	649	9	CC613918
32	81.6	6.2	733	8	BZ778636
33	81.6	6.2	925	9	CG071791
34	81.6	6.2	960	8	BZ676889
35	81.6	6.2	994	8	CC003943
36	81.6	6.2	1016	9	CL996481
37	81.6	6.2	1022	9	CL984151
38	81.4	6.2	999	8	CC385762
39	81.2	6.2	756	9	CG221653
40	81	6.2	1101	9	CNS00LT2
41	80.6	6.1	699	8	BZ996930
42	80.6	6.1	832	9	CC676387
43	80.6	6.1	839	9	CG246849
44	80.6	6.1	843	9	CG254365
45	80.6	6.1	852	8	CC385881

ALIGNMENTS

RESULT 1
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGMDO20TM ZM 0.7-1.5 KB Zea mays genomic clone ZMMMA0554D15,
genomic survey sequence.

ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.

SOURCE
ORGANISM Zea mays

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 963)

REFERENCE
AUTHORS Whicelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Unpublished (2002)

CONTRACT Other GSSs: OGMDO20TV
Contact: Cathy Whicelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whicelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..963

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZM_0.7-1.5_KB"

/note="Vector: pBCSK-; Site 1: HincII, 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 69.3%; Score 908; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 7.8e-180; Indels 0;

Matches 908; Conservative 0; Mismatches 0; Gaps 0;

404 ACCAGCATGTCTAGTGTCTACCTATGTGACCTACCAAGAAAGATTAAAAATTAACC 463
1 ACCAGCATGTCTAGTGTCTACCTATGTGACCTACCAAGAAAGATTAAAAATTAACC 60


```
QY 464 AAGTAATCATCTCGAAGAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 523
DB 61 AAGTAATCATCTCGAAGAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 120
QY 524 ACGATCCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTACAGTTTACAA 583
DB 121 ACGATCCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTACAGTTTACAA 180
QY 584 CACTAGAGCGCAATATATATGCAAGAGTATGACCTTCAGATTTTCTTTTC 643
DB 181 CACTAGAGCGCAATATATATGCAAGAGTATGACCTTCAGATTTTCTTTTC 240
QY 644 ATTCCTGTTATTTGTTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 703
DB 241 ATTCCTGTTATTTGTTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 300
QY 704 TCTTCGATTTTAAATGACTATATAAGTATTTTATATAGACGACGATGTGTAG 763
DB 301 TCTTCGATTTTAAATGACTATATAAGTATTTTATATAGACGACGATGTGTAG 360
QY 764 ATTCCTGTTCAAAATCTTCTGATTTTATAGAGTATTTGCAACCTGTCTTT 823
DB 361 ATTCCTGTTCAAAATCTTCTGATTTTATAGAGTATTTGCAACCTGTCTTT 420
QY 824 CAAGAATTTGATTTTCAAAAAAATTAAGTTTATTTCTCTTATATAAAGAAAC 883
DB 421 CAAGAATTTGATTTTCAAAAAAATTAAGTTTATTTCTCTTATATAAAGAAAC 480
QY 884 ACTTGAATAATATAGATGTCAGAGTACGCTAGAGTATTTTCCAAATAATTAATCA 943
DB 481 ACTTGAATAATATAGATGTCAGAGTACGCTAGAGTATTTTCCAAATAATTAATCA 540
QY 944 CTGTGATATATATTTTGGCAGGCCCATTAATTTTAAACGGAATCGAGCA 1003
DB 541 CTGTGATATATATTTTGGCAGGCCCATTAATTTTAAACGGAATCGAGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAAAGGAGAGAGAGAAATCG 1063
DB 601 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAAAGGAGAGAGAGAAATCG 660
QY 1064 TTTTAAGTATGTCCTCTGAGATGTCGGTGTGGCAAGATGCCATCATAGCT 1123
DB 661 TTTTAAGTATGTCCTCTGAGATGTCGGTGTGGCAAGATGCCATCATAGCT 720
QY 1124 CATAGGCTACGTAGAGTGGGAGCTCTGTGATCTCAAGAGGATCTATATAGC 1183
DB 721 CATAGGCTACGTAGAGTGGGAGCTCTGTGATCTCAAGAGGATCTATATAGC 780
QY 1184 TTGTTCAACCGTCTGTTGTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 1243
DB 781 TTGTTCAACCGTCTGTTGTCATGTCGAAGCCTTGCTATTTCTGAACCAAGAGAT 840
QY 1244 ACTTACTCCCAAAACAATCTACTATCAATGCAACTTCATGCAACGCAATATGTT 1303
DB 841 ACTTACTCCCAAAACAATCTACTATCAATGCAACTTCATGCAACGCAATATGTT 900
QY 1304 TCCTGAAC 1311
DB 901 TCCTGAAC 908

RESULT 2
CG224225 915 bp DNA linear GSS 22-ANG-2003
DEFINITION OGIAG08TV_ZM_0.7.1.5_KB Zea mays genomic clone ZMMBMA071B515,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
1 (bases 1 to 915)
WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
JOURNAL
Consortium for Maize Genomics
Unpublished (2002)
COMMENT
Other GSSs: OGIAG08TH
Contact: Cathy WhiteIaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@tigr.org
Seq primer: TP
Classes: sheared ends.
FEATURES
source
location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA071B515"
/clone_lib="ZM 0.7.1.5 KB"
/notes="Vector: pBESK-1; Site 1: HindII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 51.8%; Score 679; DB 9; Length 915;
Best Local Similarity 99.9%; Pred. No. 7.7e-132;
Matches 690; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 621 TGACCTTGAGATTTTCTTTTTCATTTCTGTTATTTGTTATGTTTATATACATTTT 680
DB 1 TGACCTTGAGATTTTCTTTTTCATTTCTGTTATTTGTTATGTTTATATACATTTT 59
QY 681 CTTCCTTACATATAGAGATTTTCTCCGATTTATATAATGACTATTAAGTCAATTTT 740
DB 60 CTTCCTTACATATAGAGATTTTCTCCGATTTATATAATGACTATTAAGTCAATTTT 119
QY 741 ATATAGAGCAGCAGATGCTAGATTCGTTCAAAAAATCTTCTGATTTTATAGAGC 800
DB 120 ATATAGAGCAGCAGATGCTAGATTCGTTCAAAAAATCTTCTGATTTTATAGAGC 179
QY 801 TAGTTGGCAACCCGTTCTTTTCAAGAAATTTTATTTTCAAAAAATTAAGTTAT 860
DB 180 TAGTTGGCAACCCGTTCTTTTCAAGAAATTTTATTTTCAAAAAATTAAGTTAT 239
QY 861 TTTCTTTTATTAATTAAGAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCTAGAAT 920
DB 240 TTTCTTTTATTAATTAAGAAACACTTAGAAAAATAGAGTTGCCAGACTAGCCTAGAAT 299
QY 921 GTTTTCCCAATTAATTAATCACTGTGTATATATTTTGGCAGCCCATTAATTAATTT 980
DB 300 GTTTTCCCAATTAATTAATCACTGTGTATATATTTTGGCAGCCCATTAATTAATTT 359
QY 981 AAACCGAACTGAATTCAGAGGAAACCAAACTGAGCTATTTCTTAAGTTATGTAAGAAG 1040
DB 360 AAACCGAACTGAATTCAGAGGAAACCAAACTGAGCTATTTCTTAAGTTATGTAAGAAG 419
QY 1041 GGAGAGAGAGAGAAAGAAATGTTTAAAGTATTTGCTCCCTGAGATGTCGGTTTGGCAA 1100
DB 420 GGAGAGAGAGAGAAAGAAATGTTTAAAGTATTTGCTCCCTGAGATGTCGGTTTGGCAA 479
QY 1101 CGATAGCCACCGTAATCACTATAGTCTATAGTGTCTTCAAGTTGGAGCTCTGTGTCT 1160
DB 480 CGATAGCCACCGTAATCACTATAGTGTCTTCAAGTTGGAGCTCTGTGTCT 539
QY 1161 ATTCACATGGCATCTATCATGCTTGTCAACCGTTCTGTTGTCATGTCGCAAGGCT 1220
DB 540 ATTCACATGGCATCTATCATGCTTGTCAACCGTTCTGTTGTCATGTCGCAAGGCT 599
```


Qy	1221	GGCCATTCTGGAACCAAGAGTACTCTCTCCCAAAACATCATCTTACTCATGCAACT	1280
Db	600	TCCTATTCTGAACCAAGAGTACTCTCTCCCAAAACATCATCTTACTCATGCAACT	659
Qy	1281	CCATGCAAAACGCAATATGTTTCTGTGAAC	1311
Db	660	CCATGCAAAACGCAATATGTTTCTGTGAAC	690
RESULT 3			
CC656939/c	687 bp	DNA	linear
LOCUS	OGMDQ20TV.ZM.0.7.1.5_KB	zea mays	genomic clone ZMBM0554D15,
DEFINITION	genomic survey sequence.		
ACCESSION	CC656939		
VERSION	CC656939.1	GI:32060231	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 687) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Remick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Cleck,R.W., Nunberg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other GSSs: OGMDQ20TM Contact: Cathy Whitelaw TIGR		
TITLE	9712 Medical Center Drive, Rockville, MD 20850, USA		
JOURNAL	Tel: 301-838-5843		
COMMENT	Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TP Class: sheared ends.		
FEATURES	location/Qualifiers		
source	1..687		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone="ZMBM0554D15"		
	/clone_1lb="ZM.0.7.1.5_KB"		
	/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		
ORIGIN			
Query Match	32.0%	Score 419.8;	DB 9; Length 687;
Best Local Similarity	99.5%;	Pred. No. 1.6e-77;	
Matches 421;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	889	GAATAATAGAGTTGGCAGACTAGAGCCCTAGATGTTTCCCAATTAATTACATCACTGTG	948
Db	687	GAATAATAGAGTTGGCAGACTAGAGCCCTAGATGTTTCCCAATTAATTACATCACTGTG	628
Qy	949	TATAATTATTTGGCCAGCCGCCATAAATATTTAAACCGAAATGTAATCGAGCAACCA	1008
Db	627	TATAATTATTTGGCCAGCCGCCATAAATATTTAAACCGAAATGTAATCGAGCAACCA	568
Qy	1009	AATCGAGCTATTTCTCTAGATTGTGTAAGGAGAGAGAGAGAAATCAATCGTTT	1068
Db	567	AATCGAGCTATTTCTCTAGATTGTGTAAGGAGAGAGAGAGAAATCAATCGTTT	508
Qy	1069	AGTCAATGTCCTGAGATGTCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG	1128
Db	507	AGTCAATGTCCTGAGATGTCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAG	448
Qy	1129	GTGCTTAGTCAAGTTCGGCAGCTTCGTGTGATCTCATAGGCACTACTAGACTGTT	1188
Db	447	GTGCTTAGTCAAGTTCGGCAGCTTCGTGTGATCTCATAGGCACTACTAGACTGTT	388

OY	1189	CAACGCTTCGTCTTGTTCATATCGTCCAAAGCCTTGCTCATTTCTGAACCAAAGGATACCTA	1248
Db	387	CAACGGTCGTCTTGTTCATATCGTCCAAAGCCTTGCTCATTTCTGAACCAAAGGATACCTA	328
OY	1249	CTCCCAACGATCCATCTTACTCATATGCACAATTCCATGCAACAGACGACATATGTTCTGT	1308
Db	327	CTCCCAACGATCCATCTTACTCATATGCACAATTCCATGCAACAGACGACATATGTTCTGT	268
OY	1309	AAC 1311 	
Db	267	AAC 265	
RESULT 4			
LOCUS	CG252571/c	715 bp	DNA linear GSS 25-AUG-2003
DEFINITION	OG4BB05TC ZM_0.7_1.5_KB zea mays genomic clone ZMMBMA0809B10,		
ACCESSION	CG252571		
VERSION	CG252571.1 GI:34154661		
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 715)		
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Renick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Numbers,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Contact: Cathy Whitelaw TIGR		
TITLE	9712 Medical Center Drive, Rockville, MD 20850, USA		
JOURNAL	Tel: 301-838-5843		
COMMENT	Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.		
FEATURES	location/Qualifiers		
Source	1..715		
	/organism="Zea mays"		
	/mol type="genomic DNA"		
	/strain="873"		
	/db xref="taxon:4577"		
	/clone="ZMMBMA0809B10"		
	/clone_id="ZM_0.7_1.5_KB"		
	/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"		
ORIGIN			
Query Match	7.3%; Score 96; DB 9; Length 715;		
Best Local Similarity	80.1%; Pred. No. 1.2e-09;		
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;			
OY	135	TCTCCCACTCGGCTTGACATTTCTTGATATGCGATGCTTCCATCGACGAGGCCCA	194
Db	356	TCAACACCCGCTTGTGTCTTCTTGATATGCGGCGTTCCTGTGATCGAGGCCCA	297
OY	195	TCAGACACCTTTCGGGACACCCATCAAGGGCTTTCGGATGGCCCCAGAGCGATTCGGG	254
Db	296	TAAAGCACCTGTGTGGAGAACCATTAAAGGCTTTTCGGATGGCCCCAGAGGACAT-CGGGG	238
OY	255	TCTGTGTGATTCAGGGGATATATGTCCCCCAATC	290
Db	237	TCTGTGTGATCCCGGGGATATCTATCCCACCAACC	202
RESULT 5			
LOCUS	CU235046	967 bp	DNA linear GSS 15-JAN-2004
DEFINITION	ZMMBB0575001r ZMMBB (HindIII) Zea mays genomic clone		

ACCESSION ZM580b0575001 3', genomic survey sequence.
VERSION CL235046
KEYWORDS CL235046.1 GI:40891729
SOURCE GSS.
ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 967)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wang,R. and Messing,J.
Sequencing of the maize genome at PGIR (2003c)
Unpublished (2003)
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 64.
Location/Qualifiers
1..967
/organism="Zea mays"
/mol_type="genomic DNA"
/cultiivar="B73"
/db_xref="taxon:4577"
/clone="ZM580b0575001"
/lab_host="E. coli DH10B"
/clone_1lib="ZM580b (HindIII)"
/note="Vector: pCUG1; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 7.0%; Score 92; DB 9; Length 967;
Best Local Similarity 78.6%; Pred. No. 8.5e-09;
Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 150 TGCACATTTTCTTGATGTCGTCGTCCTTCCATCTGACCGAGCCCATCAGACACCTTTGG 209
DB 827 TGGCCCTTTCTTGATGTCGTCGTCCTTCCATCTGACCGAGCCCATCAGACACCTTTGG 886
QY 210 GACACCATCAGAGGCGCTTCCGATGTCGTCGTCCTTCCATCTGACCGAGCCCATCAGACACCTTTGG 269
DB 887 GTGGCCCATCAGAGGCGCTTCCGATGTCGTCGTCCTTCCATCTGACCGAGCCCATCAGACACCTTTGG 946
QY 270 GGATATATGTCTCCCAACAT 289
DB 947 GGATATCTGTCTCCCAACATGT 966

RESULT 6
CG048704 814 bp DNA linear GSS 19-AUG-2003
LOCUS PU1019TB_ZM_0.6_1.0_KB Zea mays genomic clone ZM580a0611C13,
DEFINITION genomic survey sequence.
ACCESSION CG048704
VERSION CG048704.1 GI:33920884
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 814)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utecherback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)

COMMENT Other GSSs: PU1019TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source location/Qualifiers
1..814
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM580a0611C13"
/clone_1lib="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 6.7%; Score 88; DB 9; Length 814;
Best Local Similarity 62.7%; Pred. No. 5.8e-08;
Matches 153; Conservative 0; Mismatches 90; Indels 1; Gaps 1;
QY 713 TTTATTAATGACTATTAAGTCATTTTATATTAAGAGACCGCATGTCGATTCGCT 772
DB 487 TTTAAGAGACGTGTGTAGATGAGAAATATAGAAATATATCTTTAGGAATGTTA 546
QY 773 CAAAAATCTTCTGATTTTATTAAGAGCTAGTTGGCAACCTGTCTTCAAGAAAT 832
DB 547 TAAAGACACAGAAACATTTCTTAAGAGCTAGTTGTATCCCATTTTTCAGAGAT 606
QY 833 TTGATTTTTCAAA-AAAATTAAGTTTCTTCTTATTAATAAGAAACACTTAA 891
DB 607 TTGATTTTCACAGGAAATAGTTATTTCCCTTTGTAAAGTAATCATCTCGAA 666
QY 892 AAATAGAGTTCAGACTAGCCCTTAATGTTTCCCAATTAATTAACATCACTGTAT 951
DB 667 AAATAGAGTTCAGAACTAGCCCTTAAGTTATATCCCATATATCCCATATCTCT 726
QY 952 AATT 955
DB 727 ATTT 730

RESULT 7
CG414922 754 bp DNA linear GSS 08-SEP-2003
LOCUS ZM580b0290L09_r ZM580b Zea mays genomic clone ZM580b0290L09 3',
DEFINITION genomic survey sequence.
ACCESSION CG414922
VERSION CG414922.1 GI:34505144
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 754)
Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wang,R.
Sequencing of the maize genome
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers

URES source	Location/Qualifiers
	1. .754

ORIGIN

Matches 116; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Db 620 TTTCTAATAAATTAAATTCATTTTCTTGAAAAATAGAAATCACAATAGAAAAATAA 561

Db 560 GTTCCTAACTACTTAAATAATTCCTTCGTCATTACAGT 518

genomic survey sequence,
CC439901

SOURCE	Zea mays
ORGANISM	Zea mays

REFERENCE
1 (bases 1 to 950)
clade; Panicoideae; Andropogoneae; Zea.

COMMENT
Other_GSSs: PUHRV15TD
Contact: Cathy Whitelaw

Class: sheared ends.
Location/Qualifiers

COT selected genomic DNA library

Matches 126; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

Db 182 GAATAATTAGTTTATTCTCTTGAGAAATATAATCACTTGAGAAATATAGTTTCCA 123

Db 122 ACTAGCCCTAAAGTATTTTATATAAATAGAGAAAAAATTGCGTCTATTTTCATTG 69

CC384247			
LOCUS	652 bp	DNA	linear
CC384247			GSS 19-MAY-2003

ORGANISM Zea mays

AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and

CONTACT	unpublished (2003)
COMMENT	other_GSSs: PUHOC67TD
	Contact: Cathy White]aw

```

FEATURES
      Class: sheared ends.
      Location/Qualifiers

```

Cot selected genomic DNA library"

Query Match	6.6%;	Score 86.6;	DB 8;	Length 652;
Best Local Similarity	63.9%;	Pred. No. 1.1e-07;		

Db
341 TAATTATGTTAGTACTACTACATCTTATCGACATAATATGTAAGTAAATGCTA 400

401 AGAATCTATTATTAAATGTTAAGAGCTGATTTGGGAACATAATTTCTAAGAGATTTT 460

[illegible]

Db 461 TATTTTTCAGAGTAATAGTTTATTTTCATGAGGAAATAGAAATTAATGGAAAA 520
QY 894 ATAGATTGCCAGACTAGCCCTAGA 918
Db 521 ATTAGTTTCCAAACTAGCCCTAAA 545

RESULT 10
CC400575/c
LOCUS CC400575 797 bp DNA linear GSS 19-MAY-2003
DEFINITION PUHLU61TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM87A480L01,
genomic survey sequence.
ACCESSION CC400575
VERSION CC400575.1 GI:30880665
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 797)
WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHLU61TB
Contact: Cathy WhiteIaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source location/Qualifiers
1..797
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM87A480L01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.6%; Score 86.6; DB 8; Length 797;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;
Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCATGCGTAGATTCGTT 772
Db 577 TTTAGAGAGCTGGTAGAGATGGAAGATATAGAAAATATCTTTTAGAGATGTTA 518
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTCAAGAAAT 832
Db 517 TAAAGACAGAAAATTTTATAGAGCTAGTTGGTAATCCCATTTTTTCAAGAGAT 458
QY 833 TTGATTTTTCAAA-AAAATTAGTTATTTTCTTTATAAATGAAAACTAGAA 891
Db 457 TTCAATCTCACAAGGAAAATGTTATTTCCCTTTGAAAATGTAATCATCTGGAA 398
QY 892 AAATAGATTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTAT 951
Db 397 AAATAGATTCTTAAACTAGCCCTTAGAGTTATATCCGATTAATCCCTATCTCATCTCT 338
QY 952 AATTATTTGGCCAGCCCAATTAATTAATTTTAAAC 984
Db 337 ATTCAAAATTCACCTTGGCAAAATAGTGTATC 305

RESULT 11

CC400574
LOCUS CC400574 820 bp DNA linear GSS 20-AUG-2003
DEFINITION PUHLU61TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM87A480L01,
genomic survey sequence.
ACCESSION CC400574
VERSION CC400574.1 GI:30880664
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 820)
WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHLU61TD
Contact: Cathy WhiteIaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source location/Qualifiers
1..820
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM87A480L01"
/clone_lib="ZM_0.6_1.0_KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.6%; Score 86.6; DB 8; Length 820;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;
Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCATGCGTAGATTCGTT 772
Db 374 TTTAGAGAGCTGGTAGAGATGGAAGATATAGAAAATATCTTTTAGAGATGTTA 433
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTCAAGAAAT 832
Db 434 TAAAGACAGAAAATTTTATAGAGCTAGTTGGTAATCCCATTTTTTCAAGAGAT 493
QY 833 TTGATTTTTCAAA-AAAATTAGTTTATTTCTTTATAAATGAAAACTAGAA 891
Db 494 TTCAATCTCACAAGGAAAATAGTTATTTCCCTTTGTAATAGTAATCACTCGGA 553
QY 892 AAATAGATTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGTAT 951
Db 554 AAATAGATTCTTCAAACTAGCCCTTAGAGTTATATCCGATTAATCCCTATCTCATCTCT 613
QY 952 AATTATTTGGCCAGCCCAATTAATTAATTTTAAAC 984
Db 614 ATTCAAAATTCACCTTCAAAATAGTGTATC 646

RESULT 12
CG103452
LOCUS CG103452 471 bp DNA linear GSS 20-AUG-2003
DEFINITION PUBB19TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZM87A062D13,
genomic survey sequence.
ACCESSION CG103452
VERSION CG103452.1 GI:33985746
KEYWORDS GSS.
SOURCE Zea mays

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 471)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSSs: PUB019TD
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 Source
 1..471
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMB7A0628D13"
 /note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high Cor selected genomic DNA library"

ORIGIN
 Query Match 6.5%; Score 85.4; DB 9; Length 471;
 Best Local Similarity 69.5%; Pred. No. 2e-07;
 Matches 116; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 789 TTTTAAAGACGTAGTTGGCAACCGTCTTCAAGAAATTTGATTTTCAAAA 848
 |||||
 DB 216 TTTTCTTAGAGCTAGTTGGAAATCCCATTTTCCAGGATTTACATTTTTCAGGG 275
 |||||
 QY 849 AAATAGTATTTCTCTTATTAATAAGAAACACTTGAATAATAGATGCCAC 908
 |||||
 DB 276 AAATAGTATTTCTCTTGAATAATAGAAATCATGCGGAAATAGTGTCCCAAC 335
 |||||
 QY 909 TAGCCCTAGATGTTTCCCATTAATTAATCAATCAGTGTATTAAT 955
 |||||
 DB 336 TAGCGCTTAGACTATTTCTTAATAAAGTAGATCAATGAATAAGATT 382
 |||||

RESULT 13
 CC082135/ 765 bp DNA linear GSS 20-AUG-2003
 LOCUS PUF0X12TD ZM 0.6_1.0_KB Zea mays genomic clone ZMMB7A0712B24,
 DEFINITION genomic survey sequence.
 ACCESSION CC082135
 VERSION CC082135.1 GI:33964429
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 765)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other_GSSs: PUF0X12TD
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 781)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGUCG53TH
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 Source
 1..781
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMB7A0628D13"
 /note="Vector: pCR4-TOP0; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN
 Query Match 6.5%; Score 85.2; DB 9; Length 781;
 Best Local Similarity 71.8%; Pred. No. 2.2e-07;
 Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 787 ATTTTAAAGACGTAGTTGGCAACCGTCTTCAAGAAATTTGATTTTCAAA 846
 |||||
 DB 255 ATCTTCTATGACCTAGTTGGCAACCTATTT-CTCAGGATTTTATTTTTCAT 197
 |||||
 QY 847 AAAAATAGTATTTCTCTTATTAATAAGAAACACTTGAATAATAGATGCCAG 906
 |||||
 DB 196 GAAATATAGTTATTTCTCTTGAATAATATATATCATTTGAAATAATAGTTCCA 137
 |||||
 QY 907 ACTAGCCCTAGAAATGTTTCCCATTAATTAATCAATCAGTGTATTAATTTG 960
 |||||
 DB 136 ACTAGCCCTAAGATTTTATTAATAATAGAGAAATTTGCTATTTTCATTG 83
 |||||

RESULT 14
 CC630219 781 bp DNA linear GSS 19-JUN-2003
 LOCUS OGUCG53TH ZM 0.7_1.5_KB Zea mays genomic clone ZMMB7A0406J09,
 DEFINITION genomic survey sequence.
 ACCESSION CC630219
 VERSION CC630219.1 GI:32003354
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 781)
 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGUCG53TH
 Contact: Cathy WhiteLaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteLaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

Db 181 CTGACCGAGCCCATGACACACCTTTCGGGACACCCATCAAGGCGCTTTCGATGGCCCA 240
 QY 241 CGAGAGTATCGGGTGGTGGTATCCAGGGGATATATGTCCCAAAATGTGTAACCTTAA 300
 Db 241 CGAGAGTATCGGGTGGTGGTATCCAGGGGATATATGTCCCAAAATGTGTAACCTTAA 300
 QY 301 TTATTTCTTTAGATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGGTGA 360
 Db 301 TTATTTCTTTAGATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGGTGA 360
 QY 361 GGGCTTCAGATAGATTTTCGTTAGGGCCAGAAATCGAGAGACCAAGCCATGTCTAGTG 420
 Db 361 GGGCTTCAGATAGATTTTCGTTAGGGCCAGAAATCGAGAGACCAAGCCATGTCTAGTG 420
 QY 421 TCACATTTGGGCTACCCAGAACATTTAAATTTAAATACCAAGTAACTTAATCCACT 480
 Db 421 TCACATTTGGGCTACCCAGAACATTTAAATTTAAATACCAAGTAACTTAATCCACT 480
 QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAATAA 540
 Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAATAA 540
 QY 541 CAAGCATTTTGCAGAAAGACAAATTTATGTACAGTTTCAAAACATCTAAGCCGACAA 600
 Db 541 CAAGCATTTTGCAGAAAGACAAATTTATGTACAGTTTCAAAACATCTAAGCCGACAA 600
 QY 601 TTATATCGAAAGGTAAAGCTATGACGTTGAGATTTTCTTTGATTTCTTTGATTTGTT 660
 Db 601 TTATATCGAAAGGTAAAGCTATGACGTTGAGATTTTCTTTGATTTCTTTGATTTGTT 660
 QY 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTAAATA 720
 Db 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGTATTTCTTCCGATTTTAAATA 720
 QY 721 ATGACTATTAAGCATTTTATATTAAGAGACGAGATGTCTAGATTTCTGTTCAAAAATC 780
 Db 721 ATGACTATTAAGCATTTTATATTAAGAGACGAGATGTCTAGATTTCTGTTCAAAAATC 780
 QY 781 TTTCGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 Db 781 TTTCGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATTAAGAAACATCTAGAAAAATAGAGT 900
 Db 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATTAAGAAACATCTAGAAAAATAGAGT 900
 QY 901 TGGCAACTAGCCCTAGATGTTTTCCTCAATTAATTAACAATCTGTGTATTAATTTTG 960
 Db 901 TGGCAACTAGCCCTAGATGTTTTCCTCAATTAATTAACAATCTGTGTATTAATTTTG 960
 QY 961 GCCAGCCCAATAAATTTTAAACGGAACCTGAATCGAGCAACCAATCGAGCTAT 1020
 Db 961 GCCAGCCCAATAAATTTTAAACGGAACCTGAATCGAGCAACCAATCGAGCTAT 1020
 QY 1021 TTCTCAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAAGTCAATGTGCC 1080
 Db 1021 TTCTCAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAAGTCAATGTGCC 1080
 QY 1081 TGAAGATGCGGGTGGGCAACATAGCCACCTGTAATCAATAGCTTGAAGTGGCTAGTCA 1140
 Db 1081 TGAAGATGCGGGTGGGCAACATAGCCACCTGTAATCAATAGCTTGAAGTGGCTAGTCA 1140
 QY 1141 GGTTCGGAGCTCTCGTGTCTCATCTCAATGGCATATCTACATGCTTTTCAACCGTTGCTC 1200
 Db 1141 GGTTCGGAGCTCTCGTGTCTCATCTCAATGGCATATCTACATGCTTTTCAACCGTTGCTC 1200
 QY 1201 TTGTTTCATCTGTCAGAGCTTGGCTTATCTGAACCAAGAGATACCTACTCCCAACAT 1260
 Db 1201 TTGTTTCATCTGTCAGAGCTTGGCTTATCTGAACCAAGAGATACCTACTCCCAACAT 1260
 QY 1261 CCATCTTAATCTAGCAACTTCCATGCAACAGCAATATGTTTCTGTAACAGATCTATT 1320
 Db 1261 CCATCTTAATCTAGCAACTTCCATGCAACAGCAATATGTTTCTGTAACAGATCTATT 1320

QY 1321 AAAGATCAACAAGCTAGGCTTCCCGCTAGCTTCCCTCTCCCTGCGGATCTTTT 1380
 Db 1321 AAAGATCAACAAGCTAGGCTTCCCGCTAGCTTCCCTCTCCCTGCGGATCTTTT 1380
 QY 1381 CGTCCAGCACCATG 1394
 Db 1381 CGTCCAGCACCATG 1394
 RESULT 2
 BD062177
 LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
 DEFINITION Male tIssue-preferred regulatory region and method of using same.
 ACCESSION BD062177.1 GI:22607782
 VERSION BD062177.1
 KEYWORDS JP 2001520523-A/2.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 1394)
 REFERENCES
 1 (bases 1 to 1394)
 Kenda11,T.L.
 Male tIssue-preferred regulatory region and method of using same
 Patent: JP 2001520523-A 2 30-OCT-2001;
 PIONEER HI BRED INTERNATIONAL INC
 PN JP 2001520523-A/2
 PD 30-OCT-2001
 PF 19-JUN-1998 JP 1999504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAT,GARY A HUFPMAN,
 PI TIMMY L KENDALL
 PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
 ,C07K14/34,C12Q1/68,
 PC A01H5/00
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key location/Qualifiers.
 FEATURES
 source location/Qualifiers
 1..1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1394; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGATGTCCTATGAAAAAGATGATCAATGTCATATCGGTTCTTAGGGTCC 60
 Db 1 CGATGTCCTATGAAAAAGATGATCAATGTCATATCGGTTCTTAGGGTCC 60
 QY 61 CTTCTTGCCTTATTAAGTGAATCGGGGTTCAAAAAACTTTCACAGGGTGCATGAT 120
 Db 61 CTTCTTGCCTTATTAAGTGAATCGGGGTTCAAAAAACTTTCACAGGGTGCATGAT 120
 QY 121 CTCATGTTCCACTTCTCCACCTCGCGTTGCACTTTCTTGAAGTGGGTGCCAT 180
 Db 121 CTCATGTTCCACTTCTCCACCTCGCGTTGCACTTTCTTGAAGTGGGTGCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACTTTCGGGACACCCATCAAGGCTTTCCGATGGCCCA 240
 Db 181 CTGACCGAGGCCCATCAGACACTTTCGGGACACCCATCAAGGCTTTCCGATGGCCCA 240
 QY 241 CGAGAGTATCGGGTCTGTGTGATCCAGGGGATATATGTCCCAAAATCGTCACTTAA 300
 Db 241 CGAGAGTATCGGGTCTGTGTGATCCAGGGGATATATGTCCCAAAATCGTCACTTAA 300
 QY 301 TTATTTCTTTAGATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGGTGA 360
 Db 301 TTATTTCTTTAGATATTTTAAATTTTGGAAAAATACAAATCTTATCTTTGGTGA 360

QY 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCCGAGAAAATGCGAGAGACCGACGATGCTTAGTG 420
DB 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCCGAGAAAATGCGAGAGACCGACGATGCTTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAATCAATCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAATCAATCACT 480
QY 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
DB 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
QY 541 CAAGCATATTTCCAAAAGAGCAAAATTAATGTTACAGTTTAAACATCTAGACGACAA 600
DB 541 CAAGCATATTTCCAAAAGAGCAAAATTAATGTTACAGTTTAAACATCTAGACGACAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGAGTATTTTCCGATTTTATAAA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGAGTATTTTCCGATTTTATAAA 720
QY 721 ATGACATTAAGTCAATTTTATATTAAGAGACCGCATGCTAGATTTCTGTTCAAAAATC 780
DB 721 ATGACATTAAGTCAATTTTATATTAAGAGACCGCATGCTAGATTTCTGTTCAAAAATC 780
QY 781 TTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
DB 781 TTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTATTTTCTTTTAAAAATAGAAAACCTTAGAAAAATAGAGT 900
DB 841 TTCAAAAAAATAGTTATTTTCTTTTAAAAATAGAAAACCTTAGAAAAATAGAGT 900
QY 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGCTGTAATTTATTTG 960
DB 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGCTGTAATTTATTTG 960
QY 961 GCCAGCCCATTAATTAATTAACCGAACTGAAATCGAGCAAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCATTAATTAATTAACCGAACTGAAATCGAGCAAAACCAATCTGAGCTAT 1020
QY 1021 TTCTAGATTAGTAAAAAGAGAGAGAGAGAGAAATCACTTTTAAGTCAATTTGCC 1080
DB 1021 TTCTAGATTAGTAAAAAGAGAGAGAGAGAGAAATCACTTTTAAGTCAATTTGCC 1080
QY 1081 TGAAGATGCGGTTTGGCAACGATAGCCGTAATCATAGCTAGAGGCTAGGCTCA 1140
DB 1081 TGAAGATGCGGTTTGGCAACGATAGCCGTAATCATAGCTAGAGGCTAGGCTCA 1140
QY 1141 GGTTCGGAGAGCTCTCGTGTATCTCATAGGCAATGCTTGTTCACACCGTTCGTC 1200
DB 1141 GGTTCGGAGAGCTCTCGTGTATCTCATAGGCAATGCTTGTTCACACCGTTCGTC 1200
QY 1201 TTGTTTCATGTCGAAGCCCTTGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
DB 1201 TTGTTTCATGTCGAAGCCCTTGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
QY 1261 CCATCTTATCATGCACTTCCATGCAAGACGCAATATGTTTCTGAAACAGATCTATT 1320
DB 1261 CCATCTTATCATGCACTTCCATGCAAGACGCAATATGTTTCTGAAACAGATCTATT 1320
QY 1321 AAAGATCAACAAGAGCTAGGCTTCCCGGCTAGGCTTCCCTCTCTGCGGATCTTTT 1380
DB 1321 AAAGATCAACAAGAGCTAGGCTTCCCGGCTAGGCTTCCCTCTCTGCGGATCTTTT 1380
QY 1381 CGTCAACCAACATG 1394
DB 1381 CGTCAACCAACATG 1394

RESULT 3
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS AX224394
DEFINITION Sequence 1 from Patent WO0160997.
ACCESSION AX224394
VERSION AX224394.1 GI:15554636
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGTGTCTCATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAGGGTCC 60
DB 1 CCATGGTGTCTCATGAAAAAGATGAGTCAATGTCCTATATCCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTACAAAAAATTCCAGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAATGACTGAATCGGGGTTACAAAAAATTCCAGGGTGATGAT 120
QY 121 CTCGATGTTCCACTTTCACACCTCGCGTTGCACTTTCTTGATGTCGGTGTCCAT 180
DB 121 CTCGATGTTCCACTTTCACACCTCGCGTTGCACTTTCTTGATGTCGGTGTCCAT 180
QY 181 CTGACCGAGGCCCATATGACACCTTTGCGGACACCCATATAGGGCTTTGCGATGCCCA 240
DB 181 CTGACCGAGGCCCATATGACACCTTTGCGGACACCCATATAGGGCTTTGCGATGCCCA 240
QY 241 CGAGAGTATCGGGTCGTGATGATCCAGGGGATATATGTCCCAATCGTCACTATAT 300
DB 241 CGAGAGTATCGGGTCGTGATGATCCAGGGGATATATGTCCCAATCGTCACTATAT 300
QY 301 TTATTAATCTTATGATATTTATTTTGGAAAAATAACAATTAATCTTTGTGTA 360
DB 301 TTATTAATCTTATGATATTTATTTTGGAAAAATAACAATTAATCTTTGTGTA 360
QY 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCCGAGAAAATGCGAGAGACCGACGATGCTTAGTG 420
DB 361 GGGCCCTCAGATAGATTTTGGCTTAGGGCCCGAGAAAATGCGAGAGACCGACGATGCTTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAATCAATCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATAACCAAGTAATCAATCACT 480
QY 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
DB 481 CGAAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
QY 541 CAAGCATATTTCCAAAAGAGCAAAATTAATGTTACAGTTTAAACATCTAGACGACAA 600
DB 541 CAAGCATATTTCCAAAAGAGCAAAATTAATGTTACAGTTTAAACATCTAGACGACAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660

QY 661 ATGTGTTTATATACATTTCTCTCTCAATAGAGATTTCTCCGATTTTATAA 720
| | | | |
Db 661 ATGTGTTTATATACATTTCTCTCTCAATAGAGATTTCTCCGATTTTATAA 720
QY 721 ATGACTATTAAGTCATTTTATTAAGAGACGATGTCGTAGATTCTCGTCAAAAATC 780
| | | | |
Db 721 ATGACTATTAAGTCATTTTATTAAGAGACGATGTCGTAGATTCTCGTCAAAAATC 780
QY 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTT 840
| | | | |
Db 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGT 900
| | | | |
Db 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGT 900
QY 901 TGGCAGACTAGCCCTTGAATGTTTTCCCAATAAATTACATCACTGTGTATATTTTGG 960
| | | | |
Db 901 TGGCAGACTAGCCCTTGAATGTTTTCCCAATAAATTACATCACTGTGTATATTTTGG 960
QY 961 GGCAGCCCATTAATTAATTAACGGAATCGAATCGAGCGAAACCAATCTGAGCTAT 1020
| | | | |
Db 961 GGCAGCCCATTAATTAATTAACGGAATCGAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGAGAGAGAGAGAGAAATCAATTAAAGTATGTCCTCC 1080
| | | | |
Db 1021 TTCTCTAGATTAGTAAAAAGAGAGAGAGAGAGAAATCAATTAAAGTATGTCCTCC 1080
QY 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTAATCATAGTCTATAGTGTCCATCTCA 1140
| | | | |
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCCACCGTAATCATAGTCTATAGTGTCCATCTCA 1140
QY 1141 GGTGCGGAGCTCTGTCATCTCATGCGCATATACATGCTGTTGAAACCGTTCCGTC 1200
| | | | |
Db 1141 GGTGCGGAGCTCTGTCATCTCATGCGCATATACATGCTGTTGAAACCGTTCCGTC 1200
QY 1201 TTGTTCCATCTGTCGAACGCTTGGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
| | | | |
Db 1201 TTGTTCCATCTGTCGAACGCTTGGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCATGCAACAGCATATGTTTCTGTAAAC 1311
| | | | |
Db 1261 CCATCTTACTCATGCAACTTCATGCAACAGCATATGTTTCTGTAAAC 1311
RESULT 4
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS BD062176 Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062176 Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176.1 GI:22607781
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen M.C., Fox T.W., Garnat C.W., Huffman G.A. and Kendall T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL PIONEER HI BRAD INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAT, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;

FEATURES FH Key Location/Qualifiers.
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTATAGAAAAAGATAGTACATAGTGTCTATATCGCTTTCTTAGAGTCC 60
| | | | |
Db 1 CCATGCTGCTATAGAAAAAGATAGTACATAGTGTCTATATCGCTTTCTTAGAGTCC 60
QY 61 CTTCTTGCCCTTATTAAGCTAGCTGAATCGGGTTTCAAAAACTTCAAGGGTGCATGAT 120
| | | | |
Db 61 CTTCTTGCCCTTATTAAGCTAGCTGAATCGGGTTTCAAAAACTTCAAGGGTGCATGAT 120
QY 121 CTCATATGTTCCACTTCTCCACCTTGGTGAACATTTCTTGAATGTGGTGGTCCCAT 180
| | | | |
Db 121 CTCATATGTTCCACTTCTCCACCTTGGTGAACATTTCTTGAATGTGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCTTTCCGATGGCCA 240
| | | | |
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCTTTCCGATGGCCA 240
QY 241 CGAGACGATCGGGTCTGTGTGATCCAGGGGATATATGTCCGCCAATCGTCACTATA 300
| | | | |
Db 241 CGAGACGATCGGGTCTGTGTGATCCAGGGGATATATGTCCGCCAATCGTCACTATA 300
QY 301 TTATATCTTTAGATATATTTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
| | | | |
Db 301 TTATATCTTTAGATATATTTTATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
QY 361 GGGGCTCAGCATATGTTTTCGTTAGGGCCCAAGAAATCGAGACCAACCATGTCTAGTG 420
| | | | |
Db 361 GGGGCTCAGCATATGTTTTCGTTAGGGCCCAAGAAATCGAGACCAACCATGTCTAGTG 420
QY 421 TCCATATTTGGACATACCCAGAACAAAGATTTAAAAAATTAACAAAGTAACTTAATCCA 480
| | | | |
Db 421 TCCATATTTGGACATACCCAGAACAAAGATTTAAAAAATTAACAAAGTAACTTAATCCA 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCCCTTAAAAAA 540
| | | | |
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCCCTTAAAAAA 540
QY 541 CAAGCATATTTGAAAGAGACAAATTTATGTTTACAGTTTACAAACATCTTAAGAGCAAAA 600
| | | | |
Db 541 CAAGCATATTTGAAAGAGACAAATTTATGTTTACAGTTTACAAACATCTTAAGAGCAAAA 600
QY 601 TTATATCGAAAGTAAAGTATGACGTTTCAAGATTTTCTTTTCAATCTTTGTTATTTGTT 660
| | | | |
Db 601 TTATATCGAAAGTAAAGTATGACGTTTCAAGATTTTCTTTTCAATCTTTGTTATTTGTT 660
QY 661 ATGTGTTTATATACATTTCTCTCTCAATAGAGATTTCTTCGATTTTATATAA 720
| | | | |
Db 661 ATGTGTTTATATACATTTCTCTCTCAATAGAGATTTCTTCGATTTTATATAA 720
QY 721 ATGACTATTAAGTCATTTTATTAAGAGACGATGTCGTAGATTCTCGTCAAAAATC 780
| | | | |
Db 721 ATGACTATTAAGTCATTTTATTAAGAGACGATGTCGTAGATTCTCGTCAAAAATC 780
QY 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTT 840
| | | | |
Db 781 TTTTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGT 900
| | | | |
Db 841 TTCAAAAAAATTAAGTTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGT 900
QY 901 TGGCAGACTAGCCCTTGAATGTTTTCCCAATAAATTACATCACTGTGTATATTTTGG 960
| | | | |

Db 901 TGGCAGCTAGCCCTAGAAATGTTTTCCCAATTAATTCATCTAGTATTAATTTG 960
Qy 961 GCCAGCCCATTAATTTTAAACCGAATGGAATGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCATTAATTTTAAACCGAATGGAATGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTGATTTAGTAAAGGAGAGAGAGAGAAATCACTTTTAAGTCTTGTCC 1080
Db 1021 TTCTCTGATTTAGTAAAGGAGAGAGAGAGAAATCACTTTTAAGTCTTGTCC 1080
Qy 1081 TGAGATGCGGTTTGGCAAAGATAGCCAGTAACTAGTCTAGTGGCTAGCTCA 1140
Db 1081 TGAGATGCGGTTTGGCAAAGATAGCCAGTAACTAGTCTAGTGGCTAGCTCA 1140
Qy 1141 GGTTCGGCAGCTCTCGTGTCACTCAGATGCGATCTACTAGCTTGTCAACGTTGCTC 1200
Db 1141 GGTTCGGCAGCTCTCGTGTCACTCAGATGCGATCTACTAGCTTGTCAACGTTGCTC 1200
Qy 1201 TTGTTCCATCTGCTCAAGCCTTGGCTATTCTGAACCAAGAGATCTACTCTCAACAT 1260
Db 1201 TTGTTCCATCTGCTCAAGCCTTGGCTATTCTGAACCAAGAGATCTACTCTCAACAT 1260
Qy 1261 CCATCTTACTCAGCACTTCCATGCAACACGACATATGTTTCTGAAAC 1311
Db 1261 CCATCTTACTCAGCACTTCCATGCAACACGACATATGTTTCTGAAAC 1311

RESULT 5
LOCUS AF360356 3343 bp DNA linear PLN 12-MAY-2001
DEFINITION Zea mays male fertility protein (Me45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimmel,M.R. and Albertsen,M.C.
TITLE Cloning of Me45, a gene required for male fertility from Zea mays
JOURNAL 2 (bases 1 to 3343)
REFERENCE 2 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimmel,M.R. and Albertsen,M.C.
TITLE Direct Submision
JOURNAL Submitted (13-MAR-2001) Trait and Technology Development, Pioneer HI-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004, USA
FEATURES
source Location/Qualifiers
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Me45"
join(<1392..1766,1898..2182,2280..2447,2534..>2942)
/gene="Me45"
/product="male fertility protein"
join(1392..1766,1898..2182,2280..2447,2534..2942)
/gene="Me45"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRLNQMRGRGDIVQYPLHFFAALALALVADPFGSLAEV
DVRPVKHELAPYGEVMGSPRDNASRLRLRLEFVGSEVPSIEFDLQKRGYAGLA
DGRVPMGEAGMETFAVMNPMSEBVCANGVSTRKHEKEBFCRGIRAFGE
TGEIYVADAYGMVVGSGVASVAREADSDIRPANDLDVHRMGVSVEFDTSRY
SRDHLNILLBEGTGRLLKYPDTSGVHVLLKGLVFPNGVOJSEHQPLRLSETTNG

ORIGIN

RIMRYLBEPRAGEVEVFAPIBGFPPNVNSNRGOFWVAIDCCTPBOEVFAPRPLR
TLPYKPLSLKVLTKMAARRMHTVLLLBEGVVEVLDRGHEVMKLVSEVREVRK
LWIGVAHNHNIATIPYPLED"

Query Match 94.0%; Score 1311; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGCTGCTCTATGAAAAAGATGATCAATGTCTATATCCGTTTCTTGGGCTC 60
Db 1 CCATGCTGCTCTATGAAAAAGATGATCAATGTCTATATCCGTTTCTTGGGCTC 60
Qy 61 CTTCCTTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCCACGGGTCAAT 120
Db 61 CTTCCTTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCCACGGGTCAAT 120
Qy 121 CTCGATGTTCACTTCTCCACCTCGGCTGTGCAATTTCTTGGATGTGGGTGGTCCAT 180
Db 121 CTCGATGTTCACTTCTCCACCTCGGCTGTGCAATTTCTTGGATGTGGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCCATCAAGGCTTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCCATCAAGGCTTTTCGATGGCCCA 240
Qy 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCCCCAATGTCACTTAT 300
Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCCCCAATGTCACTTAT 300
Qy 301 TTATTAATCTTATGATATATTAATTTTGGAAAAATTAACAACCTTATCTTGTGTA 360
Db 301 TTATTAATCTTATGATATATTAATTTTGGAAAAATTAACAACCTTATCTTGTGTA 360
Qy 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCATGTCTAGTG 420
Db 361 GGGGCTCAGCATGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCATGTCTAGTG 420
Qy 421 TCCACTATTGGCAGTACCAAGAACAAATTTAAAAAATTAACCAAGTATCTATCTCACT 480
Db 421 TCCACTATTGGCAGTACCAAGAACAAATTTAAAAAATTAACCAAGTATCTATCTCACT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTAACACAGATCCTCTTAAGAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTAACACAGATCCTCTTAAGAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATATGTTCAGTTTCAAAACATCTAAGAGCGACAAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATATGTTCAGTTTCAAAACATCTAAGAGCGACAAA 600
Qy 601 TTATATCGAAAGGTAAAGTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGTATGACGTTCAAGATTTTCTTTTCATTTCTGTTATTTGTT 660
Qy 661 ATTTGTTTATATACATTTTCTTCTTCAATAGTATTTCTTCGATTTTATTA 720
Db 661 ATTTGTTTATATACATTTTCTTCTTCAATAGTATTTCTTCGATTTTATTA 720
Qy 721 ATGACTATAAAGTCATTTTATATAAGACAGCATGTGTGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAGACAGCATGTGTGATTTCTGTTCAAAAATC 780
Qy 781 TTTCGATTTTATTAAGACGTATGTCACACCTGTCTTTCAAAAGATTTGATTT 840
Db 781 TTTCGATTTTATTAAGACGTATGTCACACCTGTCTTTCAAAAGATTTGATTT 840
Qy 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAAGAAAACTTCAAAAAATAGAT 900
Db 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAAGAAAACTTCAAAAAATAGAT 900
Qy 901 TGGCAGTACGCGCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTGATTAATTTG 960
Db 901 TGGCAGTACGCGCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTGATTAATTTG 960

Oy	961	GCACGCCCATTAATTTATTAACCGGAATCTGAAATCGACGGAAACCAATCTGAGCTAT	1020
Db	961	GCACGCCCATTAATTTATTTAAACCGAACTGAAATCGACGGAAACCAATCTGAGCTAT	1020
Oy	1021	TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAAGAAATCAGTTTAACTCATTTGCC	1080
Db	1021	TTCTCTAATTTAGTAAAGGAGAGAGAGAGAAAGAAATCAGTTTAACTCATTTGCC	1080
Oy	1081	TGAGATGTGCGGTTTGGCAAGATAGCCACGGTATCTATAGCTCATAGTCCCTACGCTCA	1140
Db	1081	TGAGATGTGCGGTTTGGCAAGATAGCCACGGTATCTATAGCTCATAGTCCCTACGCTCA	1140
Oy	1141	GGTTGGCAGCTCTCTGTCTCATCTCACTGCGATCTATCATGCTTTGTTCAACCGTTGCTC	1200
Db	1141	GGTTGGCAGCTCTCTGTCTCATCTCACTGCGATCTATCATGCTTTGTTCAACCGTTGCTC	1200
Oy	1201	TTGTTTCCATCGTCCCAAGCTTGGCTTATTTCTGAACCAAGAGATACTACTCCCAACAAT	1260
Db	1201	TTGTTTCCATCGTCCCAAGCTTGGCTTATTTCTGAACCAAGAGATACCTACTCCCAACAAT	1260
Oy	1261	CCATCTTACTATGCAACTTCACATGCAAAACGCAATATGTTTTCCTGAAC	1311
Db	1261	CCATCTTACTATGCAACTTCACATGCAAAACGCAATATGTTTTCCTGAAC	1311

RESULT 6					
AX224402	AX224402	255 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	Sequence 9 from Patent WO0160957.				
DEFINITION	AX224402				
ACCESSION	AX224402.1	GI:15554644			
VERSION					
KEYWORDS					
SOURCE	Zea mays				
ORGANISM	Zea mays				

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.	Male tissue-preferred regulatory region and method of using same	Patent: WO 0160997-A 9 23-Aug-2001; PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES	Location/Qualifiers
source	1..255
	/organism="Zea mays"
	/mol_type="unassigned DNA"
	/db_xref="taxon:4577"

ORIGIN	
Query Match	17.2%; Score 240; DB 6; Length 255;
Beet Local Similarity	100.0%; Pred. No. 2e-113;
Matches 240; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

0y	1155	CGTGCATCTCAGACGCGATCTACGAGCTGTTCMAACCGTTCGTTGTTCCATGCGCC	1214
Db	15	CGTGCATCTCAGACGCGATCTACGAGCTGTTCMAACCGTTCGTTGTTCCATGCGCC	74
0y	1215	AAGCTTGCCATTCTGAACCAAGAGTACTCTACGCCAAACAATCATCTTTCATG	1274

Db 75 AAGCTTGTCGATTTCTGAAACCAAGAGATACCTACTCCAAACAAATCCATTACTACTG 134

Qy 1275 CCACTTCGATCCAAACAGCAGCATATGTTTCTGTAACGATCTTTAAAGTCAACAAG 1354

Db 135 CACTTCGATCCAAACAGCAGCATATGTTTCTGTAACGATCTTTAAAGTCAACAAG 194

Qy 1335 CTAGGCTTCTCCGCTAGACTCCCTCTCTCTCTGCGGATCTTTTGGTCCACCAACATG 1394

Db 195 CTAGGCTTCTCCGCTAGACTCCCTCTCTCTCTGCGGATCTTTTGGTCCACCAACATG 254

RESULT 7	
AX224396	
LOCUS	
AX224396	158 bp
	DNA
	linear
	PAT 10-SEP-2001

DEFINITION	Sequence 3 from Patent WO0160997.
ACCESSION	AX224596
VERSION	AX224596.1
KEYWORDS	GI:15554638
SOURCE	
ORGANISM	Zea mays
	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
REFERENCE	Clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	1. Albertsen, M. C., Fox, T. W., Garzaat, C. W., Huffman, G. and Kendall, T. L.
TITLE	Male tissue-preferred regulatory region and method of using same
JOURNAL	Patent: WO 0160997-A 3 23-AUG-2001;
	PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES	Location/Qualifiers
source	1. 158
	/organism="Zea mays"
	/mol_type="unassigned DNA"
	/db_xref="taxon:4577"
ORIGIN	
Query Match	8.0%; Score 111; DB 6; Length 158;
Best Local Similarity	100.0%; Pred. No. 4,8e-46;
Matches 111; Conservative	0; Mismatches 0; Gaps 0;

Oy 120 TTTGTCATGTCCTCAACCTTGCCATTCTGAA CCAAGAGAAATCCACTCCCAAAAT 1260
 Db 48 TTGTCATGTCCTCAACCTTGCCATTCTGAA CCAAGAGAAATCCACTCCCAAAAT 107
 Oy 1261 CCACTTAACTGACCAACTTCATGCAAAACGACATATGTTCCGAAAC 1311
 Db 108 CCACTTAACTGACCAACTTCATGCAAAACGACATATGTTCCGAAAC 158

RESULT	8		
LOCUS	AX224398		
DEFINITION	Sequence 5 from Patent WO016097.	50 bp	DNA
ACCESSION	AX224398		linear
VERSION	AX224398.1	GI:15554640	PAT 10-SEP-2001
KEYWORDS	.		
SOURCE	Zea mays		
ORGANISM	Zea mays		

Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1
 Albertsen, M. C., Fox, T. W., Garnaat, C. W., Huffman, G. and Kendall, T. L.
 Male tlesue-preferred regulatory region and method of using same
 Patent: WO 0160997-A 5 23-AUG-2001
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 Location/Qualifiers

```

SOURCE
1. .50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

```

Query Match	3.6%	Score 50	DB 6	Length 50
Best Local Similarity	100.0%	Pred No. 3	6e-14	
Matches 50	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Db	1229	CTGAACCAAGATGCTTACTCTCCCAACCAATCATCTTTCATGCAAC	1278	
	1	CTGAACCAAGATGCTTACTCTCCCAACCAATCATCTTTCATGCAAC	50	

RESULT 9	AX224399	LOCUS	AX224399	DEFINITION	Sequence 6 from Patent WO016097.	40 bp	DNA	linear	PAT 10-SEP-2001
ACCESSION	AX224399.1	VERSION	GI:15554641						

KEYWORDS	Zea mays
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	1 Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE	Male tLase-preferred regulatory region and method of using same
JOURNAL	Patent: WO 0160997-A 6 23-AUG-2001;
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)
SOURCE	Location/Qualifiers
1..40	
/organism="Zea mays"	
/mol_type="unassigned DNA"	
/db_xref="taxon:4577"	
ORIGIN	
Query Match	2.9%; Score 40; DB 6; Length 40;
Best Local Similarity	100.0%; Pred. No. 6,1e-09;
Matches	40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1239 AGGATACCTTACTCCCAACATCCATCTTACTGATGCAAC 1278
Db	1 AGGATACCTTACTCCCAACATCCATCTTACTGATGCAAC 40
RESULT 10	
LOCUS	AX224397 30 bp DNA linear PAT 10-SEP-2001
DEFINITION	Sequence 4 from Patent WO0160997.
ACCESSION	AX224397
VERSION	AX224397.1 GI:15554639
KEYWORDS	
SOURCE	
ORGANISM	Zea mays
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
AUTHORS	1 Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE	Male tLase-preferred regulatory region and method of using same
JOURNAL	Patent: WO 0160997-A 4 23-AUG-2001;
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)
SOURCE	Location/Qualifiers
1..30	
/organism="Zea mays"	
/mol_type="unassigned DNA"	
/db_xref="taxon:4577"	
ORIGIN	
Query Match	2.2%; Score 30; DB 6; Length 30;
Best Local Similarity	100.0%; Pred. No. 0.001;
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1179 CATGCTTGTCACCGCTCGTCTTGTCACA 1208
Db	1 CATGCTTGTCACCGCTCGTCTTGTCACA 30
RESULT 11	
LOCUS	AC149836 114817 bp DNA linear HTG 22-JUN-2004
DEFINITION	Zea mays clone ZMBB04961L7, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.
ACCESSION	AC149836
VERSION	AC149836.1 GI:49035067
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULFUP; HTGS_ACTIVEFIN.
SOURCE	Zea mays
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

```

REFERENCE      (bases 1 to 114817)
AUTHORS        Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.
TITLE          Zea mays, clone ZMMBc0496117
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 114817)
AUTHORS        Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.,
               Messing,J., Abouelellil,A., Allen,N., Anderson,M., Anderson,S.,
               Arachchi,H.M., Barina,N., Baetjen,V., Bloom,T., Boguslavsky,L.,
               Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A.,
               Cook,A., Cooke,P., Corum,B., DeArrellano,K., Diaz,J.S., Dodge,S.,
               Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P.,
               Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L.,
               Grand-Pierre,N., Hafez,N., Haeggen,D., Haags,B., Hall,J.,
               Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
               Karakas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
               Liu,G., Liu,X., Lui,A., Mabbitt,R., Maclean,C., MacDonald,P.,
               Major,J., Manning,J., Mathews,C., McCarthy,M., Meldrum,J.,
               Meneses,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
               Nguyen,T., Nicol,R., Nordu,C., O'Connor,T., O'Donnell,P.,
               O'Neill,D., Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N.,
               Rachpaka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
               Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C.,
               Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
               Talamaas,J., Tesfaye,S., Theodore,J., Tophan,K., Travetsi,M.,
               Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
               Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
COMMENT        Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome
               Research, 320 Charles Street, Cambridge, MA 02141, USA
               All repeats were identified using RepeatMasker:
               Smit,A.F.A. & Green, P. (1996-1997)
               http://ftp.genome.washington.edu/RM/RepeatMasker.html
               -----
               Genome Center
               Center: Whitehead Institute/MIT Center for Genome Research
               Center code: WIRB
               Web site: http://www-seq.wi.mit.edu
               Contact: sequence_submissions@road.mit.edu
               -----
               Bharti, A.K and Messing, J: The Plant Genome Initiative at
               Rutgers, Waksman Institute, Rutgers, The State University of New
               Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
               (http://pgr.rutgers.edu)
               Butler, E and Wing, R: Arizona Genomics Institute, Biological
               Sciences West, 448A, P.O. Box 210088, University of Arizona,
               Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
               -----
               Project Information
               Center project name: L30352
               Center clone name: 496_L_17
               -----
               * NOTE: This is a 'working draft' sequence. It currently
               * consists of 3 contigs. The true order of the pieces
               * is not known and their order in this sequence record is
               * arbitrary. Gaps between the contigs are represented as
               * runs of N, but the exact sizes of the gaps are unknown.
               * This record will be updated with the finished sequence
               * as soon as it is available and the accession number will
               * be preserved.
               *
               * 1 35156: contig of 35156 bp in length
               * 35157 35256: gap of 100 bp
               * 35257 97940: contig of 62684 bp in length
               * 97941 98040: gap of 100 bp
               * 98041 114817: contig of 16777 bp in length.
               Location/Qualifiers
                   1..114817
                       /organism="Zea mays"
                       /mol_type="genomic DNA"
                       /cultivar="B73"
                       /db_xref="taxon:4577"
                       /clone="ZMMBc0496117"
                       /clone_lib="CHORI-201 Maize B73 BAC WbOI Library"
FEATURES
source
ORIGIN

```


Query Match 1.8%; Score 25; Db 2; Length 114817;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 AAAAAAAAAATTGTTATTTCTCTT 868
|||||
Db 27886 AAAAAAAAAATTGTTATTTCTCTT 27910

RESULT 12
AE010607 10384 bp DNA linear BCT 25-MAR-2002
LOCUS Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 149 of
DEFINITION 197 of the complete genome.
ACCESSION AE010607 AE009851
VERSION AE010607.1 GI:19714578
KEYWORDS
SOURCE Fusobacterium nucleatum subsp. nucleatum ATCC 25586
ORGANISM Bacteria; Fusobacteriia; Fusobacteriales; Fusobacteriaceae;
Fusobacterium.
REFERENCE 1 (bases 1 to 10384)
AUTHORS Kapral, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,
Lykidi, A., Bhattacharya, A., Bartman, A., Gardner, W., Grechkin, G.,
Zhu, L., Vasileva, O., Chu, L., Kogan, Y., Chaga, O., Goldsman, E.,
Bernal, A., Larsen, N., D'Souza, M., Walunas, T., Pusch, G.,
Haselkorn, R., Fongstein, M., Kyrpides, N. and Overbeek, R.
TITLE Genome sequence and analysis of the oral bacterium Fusobacterium
JOURNAL nucleatum strain ATCC 25586
MEDLINE J. Bacteriol. 184 (7), 2005-2018 (2002)
PUBMED 2186394
1189109
REFERENCE 2 (bases 1 to 10384)
AUTHORS Kapral, V., Anderson, I., Ivanova, N., Reznik, G., Los, T.,
Lykidi, A., Bhattacharya, A., Bartman, A., Gardner, W., Grechkin, G.,
Zhu, L., Chu, L., Kogan, Y., Chaga, O., Goldsman, E., Bernal, A.,
Larsen, N., D'Souza, M., Walunas, T., Pusch, G., D., Haselkorn, R.,
Fongstein, M., Kyrpides, N. and Overbeek, R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park
Drive, Chicago, IL 60612, USA
FEATURES
source
1. 10384
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC
25586"
/mol_type="genomic DNA"
/strain="ATCC 25586"
/sub_species="nucleatum"
/db_xref="ATCC:25586"
/db_xref="taxon:190304"
162 .1664
/gene="FN0998"
162 .1664
/gene="FN0998"
/codon_start=1
/transl_table=11
/product="dipeptide-binding protein"
/protein_id="AAJ95194.1"
/db_xref="GI:19714579"
/translation="MKRTFLVSLFLVCGQNTKSPNGSTVVIIGQAKSLSDP
HMYNSIPDLVSRQFNLTFRKDGSIKPELASEYKNDKELDIILKGLKFDGS
ELTADVVPSFERMKDPGSSIMVEIIDYKEKNDYEIKLFLKSSSALLNLAHPIIT
SLYNNKYVAGNDLSVAPMGTAFLIANDGEKLELAFKQSPFEGAPVVEKITTPAI
PBDTSMALLETGEVDIATGMPVSTQTI EANDKDLISEPTATVEYICLANEKAPD
NKDFKALNYAIDKOSITDSITSGRGVAKSLVNVFVFGYDFEYFPNPKAKBLI
KSGGVKMSFSLVNDSPVRLQVAQIIQANLKDVGIDNMIETLEWGTYLQKTEGDFGS
AYLGGMISGTSADIVLYPLDLSKISIGFGNARVSNPEDFSEVEMARIALPBERKE
HFKNQOIIISOEDSPLIVLYKNKENIGINKRIKGFEDPTTMKFKNLKLIK"
1669 .2712
/gene="FN0999"
/gene="FN0999"
1669 .2712
/gene="FN0999"
/EC_number="3.4.11.-"

/codon_start=1
/transl_table=11
/product="deblocking aminopeptidase"
/protein_id="AAJ95195.1"
/db_xref="GI:19714580"
/translation="MILNNIDLKYLANKTEVELLNPSPVGYTHNAIEWNKELKGLGI
KVENYITKKGALAIYIKGDSNYKMIASHVDITGLAVVKKIKKASGLVTVNGFAMGS
VEGENVTIHTISGKTSGTLPIKASVHYGVVAREMPTEETMEIRLIDEDYKTEBDV
LKGLIGDPRISFETHRILDNQYIKSRVLDKLCVADQILSYIKYLKONKLEPKPDYI
IYFSYSEIIGHGVSVFPEDLDEFLAVDITGLVAGEBAHDEKVKQIIAKDSRTYPTTL
RKLDQETASKNNIKYTVGVYNNRYSDDATTAVALQGFDFYACIGPNVDATTHYERCHND
GIETVKLLIAL"
2854 .3936
/gene="FN1000"
2854 .3936
/gene="FN1000"
/EC_number="2.8.1.6"
/codon_start=1
/transl_table=11
/product="Biotin synthase"
/protein_id="AAJ95196.1"
/db_xref="GI:19714581"
/translation="WLKKNSSAGGKFNFFNFKKEDNNQAEPSNVKEFISYLNKKII
NEKYETREBAIFLSQIPNNMETLNIIFNADQIREVFCGKYPDLCITINAKSKCS
ENCKYCAQSVHKTGTVDYGLISKELALYEARENEGARSLVTSGLANGKEKL
DKLVIYKYIGHSTKLELCASHGICTKEALQKLDVAGVLTTHNLLESRRRYPNVCT
SHSYDRINTKNKAYAGLDVCSGGRIGLGETIEDRIDMALDRLALETHSVPIVLTPT
ITGTFPENNEEVPFEIITKTSITRFIMPKSLRYCGGRITKLGEHAKTGLRGINSAL
TGNFLTITGTEIDTKMKIKELGYEI"
3926 .4585
/gene="FN1001"
3926 .4585
/gene="FN1001"
/EC_number="6.3.3.3"
/codon_start=1
/transl_table=11
/product="Dethiobiotin synthetase"
/protein_id="AAJ95197.1"
/db_xref="GI:19714582"
/translation="WKFKDFYIGTDTDVGKTYVSTLYKALKKRNFOYKRIQSGCF
LKDGKLTAPDVDFLTGVGIDVDSMTVITLKEVSPHLASMBEGRILEIVKQHYE
DLKKHNSNIVLVEAGGLVPLRIDIFYIYIKLPNLVVYLVCGRVGAINTMTLNT
ALDTWGIKTLHGLVFNNNYRGOFEDDNIKVIALSKIKYLLIINKQKEISDEIEKFF
N"
4653 .6011
/gene="FN1002"
4653 .6011
/gene="FN1002"
/EC_number="2.6.1.62"
/codon_start=1
/transl_table=11
/product="adenosylmethionine-8-amino-7-oxononanoate
aminotransferase"
/protein_id="AAJ95198.1"
/db_xref="GI:19714583"
/translation="WRGVCIIMINNSLQKKDLKYVFPACQAKDFEENPPLVIKKGD
GLYIDENGNKMYCDISGMWYNIHGHCKRINRIITTEOVNNLEHVPANFTHEPAEL
GBELTKVLPKGIKNTLSPDNGSSCEIEMALKSPQHLDTGNPQTKTPISENAHYGET
IGALGVGDVDTFTETRPILIKGRKRVPIYDSKLSNDFVLEBCEIKELDLTIKN
HNEIACMTVEPNOVQAAGIKIYSAKFLAARDLTQKVIHLIDDEIANGFGGTGMFA
CEHAQIIEBDMNACIAGSSGVYPIAMLCTITDDIPNAPYADYEGKSPHSHYSGNPL
GRIALVELRIFKENIILITNEKQTVLRNKKQIFEDKSYIKDIRNLTGALGALDK
NLRLPVRRIGIYNALKKGAFLRIGNSYFWMPPYVITYEIDKMLVCKESIBEL
KI"
6737 .7558
/gene="FN1003"
6737 .7558
/gene="FN1003"
/codon_start=1
/transl_table=11
/product="Outer membrane protein P1 precursor"
/protein_id="AAJ95199.1"
/db_xref="GI:19714584"


```

/translation="MDAATGKGLSAAOIAIAIKTKKTNALAKLQORVNLSONGSG
DISCKRAGKGPOLGIVNYKNDKLANAARDSDIKANPAKAGHEHLEETDILAKOTI
GLSTFYPOYTINSKIRKDLPAITLSVGASYKADNYLVSTTANYPFNQADMDRPTTG
EHEGRDYKNGWEIAVENEYKLNKDFLISGLNANAGATASYNDEYANDEYFLTG
GIRYODESUALITASVAHFYQGAEGNFKEKYGTENQKYKKEITAVGLSVTKKF"
7577. .8143
/gene="FN1004"
7577. .8143
/codon_start=1
/translation="Transcriptional regulator, Tetr family"
/product="Transcriptional regulator, Tetr family"
/protein_id="AA195200.1"
/db_xref="GI:19714586"
/translation="MPKKVLSPKSFILDKSPFLKEGIGISISARNVCKLIDSPADI
YKISGSMKNKKEIKAKDLFEYLKRGKGIFFLDIGKGISIFABEQLFLQVDS
KDNIEGLIDELFLIRREIKKDELLIRIDKEQEBLLVCWFPAHGLSTLIATGFK
NPDEFEKYLADAPARLFYEYENYSK"
8337. .8810
/gene="FN1005"
8337. .8810
/gene="FN1005"
/codon_start=1
/translation="Hypothetical protein"
/product="Hypothetical protein"
/protein_id="AA195201.1"
/db_xref="GI:19714586"
/translation="MINQNAORFLDSQOLEVDORRERENAAAQYQEPDRDI
YHRAVAVNMEDTGHFYGPPELLMTFSPNSKKAIAIKAKKEPRIGADLDIKATYH
LIMDGLRYIVMGNTKKTGKMEAFKFEVDENIKILINKCSITCONCPFSNS"
complement(8863. .9261)
/gene="FN1006"
complement(8863. .9261)
/EC_number="2.3.1.-"
/gene="FN1006"
/codon_start=1
/translation="Acetyltransferase"
/product="Acetyltransferase"
/protein_id="AA195202.1"
/db_xref="GI:19714587"
/translation="MECKIINKDNTNVLNDLTKLNTSYAKDRKKEKYKVEKSLC
YPAVDTKNLIGFAIRAITDYNTNYLCDIIVDEYNGKIGKVLVETLLINDDLQY
RGLITDKAKKEFKFVKNEDVMQDKK"
9433. .10317
/gene="FN1007"
9433. .10317
/gene="FN1007"
/codon_start=1
/translation="unknown"
/product="unknown"
/protein_id="AA195203.1"
/db_xref="GI:19714588"
/translation="MKENNNNRKHKRDLVFLIWNILIFGLSIFLPSLNNKVDVS
DIBVYIIIGFLIISLITIKIKIKLAFSPKSNIIITKEKEIHAYKLVNNVKK
LVYRMVIFAPILFSFYVIFLIATSKIQNNREYVSLIYGKTIPTITLPCIV
IPLIYVSSQKIMLQKVDYFASKEVNNLDETESLTELTPKPKITFRKRFIMD
GSLNIFLIEDIEKVEKYKYPFYITKLLIHKNGRKKIRYAGPDENBMRKNFIV
KNSNIEGKVEYHINLPI"

```

```

ORIGIN
Query Match 1.7%; Score 24; DB 1; Length 10384;
Bseq Local Similarity 100.0%; Pred. No. 1.2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 872 AAAATAGAAAACCTTAGAAAAAT 895
DB 2719 AAAATAGAAAACCTTAGAAAAAT 2742

```

```

RESULT 13 AC113237 184509 bp DNA linear HTG 07-AUG-2002
LOCUS AC113237/c Canis familiaris clone RP81-70J12, WORKING DRAFT SEQUENCE, 3
DEFINITION

```

```

ACCESSION AC113237
VERSION AC113237.2 GI:22128672
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 184509)
AUTHORS Ahlter, N., Antonellis, A., Ayle, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Ido, J.R., Karlins, E., Laric, P.,
Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,
Margulies, E.H., Mastello, C., Maekert, B., Maestrian, S.D.,
McCloskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantioop, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsungeo, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 184509)
Green, E.D.
Direct Submission
Submitted (27-FEB-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
3 (bases 1 to 184509)
Green, E.D.
Direct Submission
Submitted (07-AUG-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
On Aug 7, 2002 this sequence version replaced gi:18958664.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
----- Project Information
Center project name: avx
Center clone name: 070J12

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phred-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 184204 bases at least Q40
 Consensus quality: 184261 bases at least Q30
 Consensus quality: 184294 bases at least Q20
 Insert size: 16200; agarose-fp
 Insert size: 184309; sum-of-contigs
 Quality coverage: 10.51x in Q20 bases; agarose-fp
 Quality coverage: 9.24x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and


```

* the accession number will be preserved.
* 1 127823: contig of 127823 bp in length
* 127824 127923: gap of unknown length
* 127924 147441: contig of 19518 bp in length
* 147442 147541: gap of unknown length
* 147542 184509: contig of 36968 bp in length.
Location/Qualifiers
1. 184509
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/clone="RP81-7012"
/clone_lib="RP81"
1. 127823
/notes="assembly_fragment"
clone_end:7
vector_side:left"
1. 3958
/notes="clone overlaps with GenBank Accession Number
AC110671 clone RP81-69D24 (center project name aw)"
127924. 147441
/notes="assembly_fragment"
129971. 184509
/notes="clone overlaps with GenBank Accession Number
AC113567 clone RP81-265M12 (center project name cio)"
147542. 184509
/notes="assembly_fragment"
clone_end:SP6
vector_side:right"

ORIGIN
Query Match 1.7%; Score 24; DB 2; Length 184509;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AAAAGGAGAGAGAGAGAGAA 1059
DB 23292 AAAAGGAGAGAGAGAGAGAA 23269

RESULT 14
AC098210/c 212621 bp DNA linear HTG 13-MAY-2003
LOCUS AC098210
DEFINITION Rattus norvegicus clone CH230-176E2, WORKING DRAFT SEQUENCE.
ACCESSION AC098210
VERSION AC098210.8 GI:30581396
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 212621)
Muzny,D.,Matre,M.,Metzker,M.,Lee,A.,Abramson,S.,Adams,C.,Alder,J.,
Allen,C.,Allen,H.,Alsbrooks,S.,Amin,A.,Anguiano,D.,
Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderson,E.,
Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Cencer,A.,
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
Davila,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,
Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,
Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Eaves,K.,
Egan,A.,Bacotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
Fraser,C.M.,Gabriel,A.,Ganta,R.,Garcia,A.,Garner,T.,Garcia,M.,
Gebregiorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,
Gunaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,
Harvey,Y.,Havlak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
Hollins,B.,Howells,S.,Huiyk,S.,Hume,J.,Idlebird,D.,Jackson,A.,
Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolivet,A.,
Karpathy,S.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Koyar,C.,
Kowis,C.,Kraft,C.L.,Lebow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,
Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
Lorenshew,L.,Louisege,H.,Lozada,R.U.,Lu,X.,Ma,J.,
Maheshwari,M.,Mahindaratne,M.,Maimoud,M.,Mallory,K.,Mangum,A.,
Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,
Mawhney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,E.,
Milosavljevic,A.,Miner,G.,Munja,E.,Montenayor,J.,Moore,S.,
Morgan,M.,Morris,K.,Morris,S.,Munidas,M.,Murphy,M.,Nair,L.,
Nankervis,C.,Neal,D.,Newton,N.,Nguyen,N.,Norris,S.,Parks,K.,
Nwackelamen,O.,Okwunonu,G.,Olatunmbaagun,A.,Pal,S.,Parks,K.,
Pasernak,S.,Paul,H.,Perez,A.,Perez,L.,Pfankuch,C.,
Plopper,F.,Poindexter,A.,Popovic,D.,Prims,E.,Pu,L.,L.,
Puazo,M.,Quiroz,J.,Rachin,E.,Reeves,K.,Regier,M.A.,Reigh,R.,
Reilly,B.,Reilly,M.,Ren,Y.,Reuter,M.,Richards,S.,Riggs,F.,
Rivers,C.,Rodey,T.,Rojas,A.,Rose,M.,Rose,R.,Rutz,S.U.,
Sanders,M.,Savery,G.,Scherer,S.,Scott,G.,Shatman,S.,Shen,H.,
Shetty,J.,Shvartsbeyn,A.,Sisson,I.,Slitter,C.D.,Smajls,D.,
Sneed,A.,Sodergren,E.,Song,X.-Z.,Sorelle,R.,Soza,J.,
Steinle,M.,Strong,R.,Sutton,A.,Svatek,A.,Tabor,P.,Taylor,C.,
Taylor,T.,Thomas,N.,Thomas,S.,Tingey,A.,Trefos,Z.,Uemari,K.,
Vales,R.,Vera,V.,Villasana,D.,Walton,L.,Walker,B.,Wang,J.,
Wang,Q.,Wang,S.,Warren,J.,Warren,R.,Wei,X.,White,F.,
Williams,G.,Wilson,R.,Wiczysk,R.,Wooden,H.,Worley,K.,
Wright,D.,Wright,R.,Wu,J.,Yakub,S.,Yen,J.,Yoon,L.,Yoon,V.,
Yu,P.,Zhang,J.,Zhou,J.,Zhou,X.,Zhao,S.,Zhao,S.,Dunn,D.,von
Niederhausern,A.,Weise,R.,Smith,D.R.,Holt,R.A.,Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission

Unpublished
2 (bases 1 to 212621)
Worley,K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212621)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23268092.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GILN
Center clone name: CH230-176E2
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 205647 bases at least Q40
Consensus quality: 207827 bases at least Q30
Consensus quality: 209298 bases at least Q20
Estimated insert size: 214875; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

```



```

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 212621: contig of 212621 bp in length.
Location/Qualifiers
1. 212621
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-176E2"
1828..2550
/note="clone_boundary
clone_end:T7
site:ECORI
end_sequence:BH359530"

ORIGIN
Query Match 1.7% Score 24; DB 2; Length 212621;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATTAGTTATTTCTCTTAT 871
1464 AAAATTAGTTATTTCTCTTAT 1441

Db 1464 AAAATTAGTTATTTCTCTTAT 1441

RESULT 15
AC110715/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-155J23, *** SEQUENCING IN PROGRESS
*** 2 unordered pieces.
AC110715
AC110715.5 GI:25072769
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 213753)
Muzny,D,Marie, Metzker,M, Lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alibrooke,S, Amin,A, Anguiano,D,
Anyalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Blewett,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Butrell,K, Calderon,E,
Cardenas,V, Carter,K, Cavazos,I, Cessari,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dedrich,D,
Delgado,O, Denson,S, Derram,C, Ding,Y, Dinh,C, Diya,K,
Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Faller,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Frazer,C,M, Gabler,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Georgakopoulos,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,M,
Gunnaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Henderson,R, Hines,S, Hladun,S,L, Hodgson,A, Hognes,M,
Hollins,B, Howells,S, Huylk,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
Kowitz,C, Kraft,C,L, Ledow,H, Levan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,M, Liu,Y, London,P, Longacre,S, Lopez,D,
Lorenshewa,L, Louieged,H, Lozano,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A,

```

```

Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mawhinney,S, McLeod,M,P, McNeill,T,Z, Meenan,E,
Migoslavjevic,A, Miner,G, Minga,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Munlaska,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwokeneme,O, Okwuonu,G, Olampinsagoon,A, Pal,S, Parks,K,
Pasternak,S, Paul,H, Perez,A, Perez,L, Frankoch,C,
Ploper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L,-L,
Puzos,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rives,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
Sanders,M, Saverly,G, Scherer,S, Scott,G, Shlaman,S, Shen,H,
Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D,
Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,K, Soea,J,
Steinle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C,
Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umanai,K,
Valas,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J,
Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F,
Williams,G, Willson,R, Wiczek,R, Woden,H, Worley,K,
Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V,
Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von
Niederhausen,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O,
Weinstock,G, and Gibbs,R,A.
Direct Submission
Unpublished
2 (bases 1 to 213753)
Worley,K,C.
Direct Submission
Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 213753)
Rat Genome Sequencing Consortium.
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSTRM
Center clone name: CH230-155J23
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 203043 bases at least Q40
Consensus quality: 204685 bases at least Q30
Consensus quality: 205672 bases at least Q20
Estimated insert size: 206221; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```


* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved

* 1 212625: contig of 212625 bp in length
* 212626 212725: gap of unknown length
* 212726 213753: contig of 1028 bp in length.
Location/Qualifiers

1. 213753

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-155J23"

1. 1535

/note="wgs_end_extension"

clone_end:Sp6"

1586. 2703

/note="wgs_end_extension"

clone_end:Sp6"

5028. 5742

/note="clone boundary"

clone_end:Sp6"

site:

end_sequence:BH364626"

33891. 35390

/note="wgs_contig"

21446. 212625

/note="wgs_contig"

Search completed: September 16, 2005, 15:25:40
Job time : 4107 secs

Query Match 1.7%; Score 24; DB 2; Length 213753;
Best Local Similarity 100.0%; Pred.No.1.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATTAGTTATTCTCTTTAT 871
DB 65390 AAAATTAGTTATTCTCTTTAT 65367

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 08:08:50 ; Search time 578.5 Seconds
(Without alignments)
14264.684 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 cccatgctgcctctcatgaaaa.....cttttcgcccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20028:*
6: geneseqn20038:*
7: geneseqn20048:*
8: geneseqn20058:*
9: geneseqn20068:*
10: geneseqn20078:*
11: geneseqn20088:*
12: geneseqn20098:*
13: geneseqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1394	100.0	1394	2	AAx07409	Aax07409 Zea mays
2	1394	100.0	1394	2	AAH76333	Aah76333 Z. mays M
3	1311	94.0	1394	2	AAx07408	Aax07408 Zea mays
4	1311	94.0	1394	5	AAH76332	Aah76332 Z. mays M
5	240	17.2	255	5	AAH76340	Aah76340 Z. mays M
6	111	8.0	158	5	AAH76334	Aah76334 Z. mays M
7	50	3.6	50	5	AAH76336	Aah76336 Z. mays M
8	40	2.9	40	5	AAH76337	Aah76337 Z. mays M
9	30	2.2	30	5	AAH76335	Aah76335 Z. mays M
10	23	1.6	849	6	ABL40437	Abi40437 Sunflower
11	21	1.5	333	5	ABV56505	Abv56505 Human pro
12	21	1.5	459	4	ABA58867	Abas58867 Human pro
13	21	1.5	459	4	AAI38581	Aai38581 Probe #72
14	21	1.5	459	4	AAK32768	Aak32768 Human bon
15	21	1.5	459	4	AAK07027	Aak07027 Human bra
16	21	1.5	459	4	ABs32487	Abes32487 Human liv
17	21	1.5	459	6	ABs07566	Abes07566 Human gen
18	21	1.5	965	3	AAa65960	Aaas65960 Pinus rad
19	21	1.5	9255	6	ABL34152	Abi34152 Human imm
20	21	1.5	9265	6	AAa63354	Aas63354 Chemical1

21	21	1.5	18817	6	ABL34494	Abi34494 Human met
22	21	1.5	18817	6	ABL70161	Abi70161 Chemical1
23	21	1.5	18817	7	ADs99755	Ades99755 Bisulphite
24	21	1.5	26493	12	ADJ12386	Adj12386 DNA fragm
25	20	1.4	260	5	ABV60726	Abv60726 Human pro
26	20	1.4	680	6	ABO18134	Abot18134 Oligonuci
27	20	1.4	680	6	ABO18135	Abot18135 Oligonuci
28	20	1.4	2523	2	AAH76376	Aah76376 Zea mays
29	20	1.4	10439	6	AAH7817	Aah7817 Alternati
30	20	1.4	11779	6	AAH42891	Aad42891 Maize sug
31	20	1.4	51198	11	ACN45000	Acn45000 Mouse sug
32	20	1.4	91071	11	ACN44004	Acn44004 Mouse gen
33	20	1.4	110000	6	ABA90193_2	ABA90193_2 of
34	20	1.4	110000	6	ABA90193_3	ABA90193_3 of
35	20	1.4	110000	6	ABQ87681_2	ABQ87681_2 of
36	20	1.4	110000	6	ABQ87681_3	ABQ87681_3 of
37	20	1.4	110000	8	ABX33717_2	ABX33717_2 of
38	20	1.4	110000	8	ABX33717_3	ABX33717_3 of
39	20	1.4	148497	12	ADO18925	Ado18925 Human bof
40	20	1.4	194534	12	ADO97481	Ado97481 Human can
41	20	1.4	337344	13	ABD32715	Abd32715 Human can
42	19	1.4	418	13	ADR65258	Adr65258 Cotton CD
43	19	1.4	468	2	AAI19065	Aai19065 Human PPA
44	19	1.4	485	4	ABa57528	Abas7528 Human pro
45	19	1.4	485	4	AAI37083	Aai37083 Probe #57

ALIGNMENTS

RESULT 1
AAx07409 standard; DNA; 1394 BP.
ID AAx07409
XX
AC AAx07409;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Me45 male tissue-preferred regulatory region.
XX
KW Me45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN W09859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR WPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Me45 male tissue-preferred regulatory region.
XX
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Me45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX Me45), which encodes a product selected from auxins, rolg and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1394; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGGCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTCTTAAGGTC 60
 DB 1 CCAATGGTCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTCTTAAGGTC 60
 QY 61 CTTCTTCGCTTATTACTGATCGGAGTTACAAAACTTCCACGGGTGATAT 120
 DB 61 CTTCTTCGCTTATTACTGATCGGAGTTACAAAACTTCCACGGGTGATAT 120
 QY 121 CTCGATGTTCACTTCTCCCACTCGGTGTCACATTTCTTGGATGTGGTGGTCCAT 180
 DB 121 CTCGATGTTCACTTCTCCCACTCGGTGTCACATTTCTTGGATGTGGTGGTCCAT 180
 QY 181 CTGACGAGGCCCATCAGACACCTTTGGGACCCCATCAAGGGCCCTTGGATGGCCCA 240
 DB 181 CTGACGAGGCCCATCAGACACCTTTGGGACCCCATCAAGGGCCCTTGGATGGCCCA 240
 QY 241 CGAGAGCTATCGGTCGTGTGTATCCAGGGATATATGTCCTCCACATGTCACCTATA 300
 DB 241 CGAGAGCTATCGGTCGTGTGTATCCAGGGATATATGTCCTCCACATGTCACCTATA 300
 QY 301 TTATATCTCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTTGTGTAT 360
 DB 301 TTATATCTCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTTGTGTAT 360
 QY 361 GGGCTCAGCATAGATTTTGGTGAAGGCCAGAAAAATGCGAGACCGCATGTCTAGTG 420
 DB 361 GGGCTCAGCATAGATTTTGGTGAAGGCCAGAAAAATGCGAGACCGCATGTCTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAATACCAAGTAACTAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAAACAAGATTTAAAAATACCAAGTAACTAATCCACT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAAATCTATTAACACAGATCCTCTAATAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAATCTATTAACACAGATCCTCTAATAAAA 540
 QY 541 CAAGCATATTTGCAAAAGAGACAATATATGTTACAGTTTCAAAACATCTAAGACGCAAA 600
 DB 541 CAAGCATATTTGCAAAAGAGACAATATATGTTACAGTTTCAAAACATCTAAGACGCAAA 600
 QY 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTTCTTTCATTTCTGTATTTTGT 660
 DB 601 TTATATCGAAAGGTAAAGTATGACGTTTCAATTTTCTTTCATTTCTGTATTTTGT 660
 QY 661 ATTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCGATTTTATAA 720
 DB 661 ATTGTTTTTATATACATTTTCTCTTCAATAGAGTATTTTCTTCGATTTTATAA 720
 QY 721 ATGACTATAAAGTCATTTTATATTAAGAGCAGCATGCTAGATTTCTGTTCAAAATC 780
 DB 721 ATGACTATAAAGTCATTTTATATTAAGAGCAGCATGCTAGATTTCTGTTCAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAAAGATTTTGAATTT 840
 DB 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTTCTTCAAAAGATTTTGAATTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTAGAAAAATGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTAGAAAAATGAGT 900
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCATATAATTAACATCACTGTATATTTTGG 960
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCATATAATTAACATCACTGTATATTTTGG 960
 QY 961 GCCAGCCCCATAATTTTAAACCGAAATCGAAATCGAAGCAAACTGAGCTAT 1020
 DB 961 GCCAGCCCCATAATTTTAAACCGAAATCGAAGCAAACTGAGCTAT 1020

DB 961 GCCAGCCCCATAATTTTAAACCGAAATCGAAATCGAAGCAAACTGAGCTAT 1020
 QY 1021 TTCTCTAGATTTGTAATAAAGGAGAGAGAGAAATCGATTTAAGTATGTCCC 1080
 DB 1021 TTCTCTAGATTTGTAATAAAGGAGAGAGAGAAATCGATTTAAGTATGTCCC 1080
 QY 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACGTAATCATAGGCTATAGGTGCTACGTCA 1140
 DB 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACGTAATCATAGGCTATAGGTGCTACGTCA 1140
 QY 1141 GGTTCGGAGCTCTCGTGTATCTCATGAGCATATACATGCTTGTTCMAACGTTGCTC 1200
 DB 1141 GGTTCGGAGCTCTCGTGTATCTCATGAGCATATACATGCTTGTTCMAACGTTGCTC 1200
 QY 1201 TTGTTCCATTCGTCGAAGCTTCCCTATTTGTGAACCAAGAGATTAATCTCCAAACAT 1260
 DB 1201 TTGTTCCATTCGTCGAAGCTTCCCTATTTGTGAACCAAGAGATTAATCTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGCAATATGTTTCTGMAACAGATCTATT 1320
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGCAATATGTTTCTGMAACAGATCTATT 1320
 QY 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGGATCTTTT 1380
 DB 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGGATCTTTT 1380
 QY 1381 CGTCCACACGCAATG 1394
 DB 1381 CGTCCACACGCAATG 1394

RESULT 2
 AAH76333
 ID AAH76333 standard; DNA; 1394 BP.
 AC AAH76333;
 XX
 AC
 XX
 AC
 XX
 DT 29-OCT-2001 (first entry)
 XX
 XX
 DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN WO20016097-A2.
 XX
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albreten MC, Fox TW, Garnaat CW, Huffman G, Kendal TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 PS Claim 4; Page 47; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (II) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC Ms45), which encodes a product selected from auxins, rolB and diapherica
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 94.0%; Score 1311; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAAAGGCTC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTATATCCGTTTCTTAAAGGCTC 60
 QY 61 CTCTCTGCTGCTTATTACTGATGAAATCGGGGTTTACAAAACTTCCAGGGGTCATGAT 120
 DB 61 CTCTCTGCTGCTTATTACTGATGAAATCGGGGTTTACAAAACTTCCAGGGGTCATGAT 120
 QY 121 CTCATGCTTCCACTTCTCCCACTCGGTTGCACTTTCTGGATGTCGGTGGTCCAT 180
 DB 121 CTCATGCTTCCACTTCTCCCACTCGGTTGCACTTTCTGGATGTCGGTGGTCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACTTTGCGGACACCCATCAAGGGCTTTGCGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACTTTGCGGACACCCATCAAGGGCTTTGCGATGGCCCA 240
 QY 241 CGAGAGCTATCGGGTGGTGGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
 DB 241 CGAGAGCTATCGGGTGGTGGATCCAGGGGATATATGTCGCCCAATCGTCACTATA 300
 QY 301 TTATTAATCTTTAGATATTAATTTTGAATAAATAAACAATTAATACCTTTTGCTA 360
 DB 301 TTATTAATCTTTAGATATTAATTTTGAATAAATAAACAATTAATACCTTTTGCTA 360
 QY 361 GGGGCTCAGATGATTTTGGCTTAAAGGCCCCAATAATGAGAGACCAAGCATGCTAAGG 420
 DB 361 GGGGCTCAGATGATTTTGGCTTAAAGGCCCCAATAATGAGAGACCAAGCATGCTAAGG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAATAAACAATTAATACCTTACCT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAACAATTAATACCTTACCT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAATAACGATCCTCTTAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAATAACGATCCTCTTAAAAA 540
 QY 541 CAAGCATTTTTCGAAGAGACAAATTAATGTAAGTTTACAAACATCTAAGAGCCGAA 600
 DB 541 CAAGCATTTTTCGAAGAGACAAATTAATGTAAGTTTACAAACATCTAAGAGCCGAA 600
 QY 601 TTATATCGAAGAGTATGCTATGAGTTCAGATTTTCTTTTCATTTCTGTTATTTTGT 660
 DB 601 TTATATCGAAGAGTATGCTATGAGTTCAGATTTTCTTTTCATTTCTGTTATTTTGT 660
 QY 661 ATTTGTTTTATATACATTTTCTCTTACATATGAGTATTTTCTTCGATTTTATAA 720
 DB 661 ATTTGTTTTATATACATTTTCTCTTACATATGAGTATTTTCTTCGATTTTATAA 720
 QY 721 ATGACATATAAGCATTTTATATTAAGACGACGATGCTGATGATCTGTTCAAAAAAC 780
 DB 721 ATGACATATAAGCATTTTATATTAAGACGACGATGCTGATGATCTGTTCAAAAAAC 780
 QY 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAAAGATTTGATTT 840
 DB 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTCAAAAGATTTGATTT 840
 QY 841 TTCAAAAAAATTAAGTTTATTTCTTTTAAATAAAGAAAACCTTAGAAAAATAGAGT 900
 DB 841 TTCAAAAAAATTAAGTTTATTTCTTTTAAATAAAGAAAACCTTAGAAAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTTCCAAATAAATTACATCATGCTGATATTAATTGG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTTCCAAATAAATTACATCATGCTGATATTAATTGG 960

QY 961 GCCAGCCCATATAATTTATTTAAACGAAACTGAATCGAGGAAACCAATCGAGCTAT 1020
 DB 961 GCCAGCCCATATAATTTATTTAAACGAAACTGAATCGAGGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATAATCATGTTTAAAGTATGCTCC 1080
 DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATAATCATGTTTAAAGTATGCTCC 1080
 QY 1081 TGAGATGTCGGTTTGGCAACATATGCCAGCTAATCATAGCTCATAGTGGCTTACGTCA 1140
 DB 1081 TGAGATGTCGGTTTGGCAACATATGCCAGCTAATCATAGCTCATAGTGGCTTACGTCA 1140
 QY 1141 GGTTCGGAGCTCTCGTGTGATCTGACATGAGCATGCTATCATGCTTTTCAACGGTTCGC 1200
 DB 1141 GGTTCGGAGCTCTCGTGTGATCTGACATGAGCATGCTATCATGCTTTTCAACGGTTCGC 1200
 QY 1201 TTGTTTCATCGTCAAGCCTTGCCTTATTTCTGAACCAAGAGATTAATCTCCAAACAT 1260
 DB 1201 TTGTTTCATCGTCAAGCCTTGCCTTATTTCTGAACCAAGAGATTAATCTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTCTGTGAC 1311
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTCTGTGAC 1311

RESULT 4

AAH76332
 ID AAH76332 standard; DNA; 1394 BP.

AAH76332;

29-OCT-2001 (first entry)

Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

Ms45; male tissue; regulatory region; transcription; male fertility;

hybrid seed; ds.

Zea mays.

WO200160997-A2.

23-AUG-2001.

13-FEB-2001; 2001WO-US004527.

15-FEB-2000; 2000US-00504487.

(PION-) PIONEER HI-BRED INT INC.

Albertain MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

WPI; 2001-514772/56.

A male tissue-preferred regulatory region comprising nucleotide sequences essential for initiating transcription of the Ms45 gene useful for mediating fertility in a male plant.

Claim 4; Page 46; 50pp; English.

The invention provides a male tissue-preferred regulatory region (i) comprising nucleotide sequences essential for initiating transcription of the Ms45 gene. A method of mediating male fertility in a plant is provided that involves introducing an expression vector comprising a promoter operably linked to (i) into a plant where the exogenous gene impacts male fertility of the plant and (ii) controls expression of the exogenous gene. A method of producing hybrid seeds is also provided. The present sequence represents a nucleic acid sequence encoding an Ms45 male tissue preferred regulatory region from Z. mays

Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 94.0%; Score 1311; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGATGTCCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAAGGTC 60
 Db 1 CCATGATGTCCTATGAAAAAGATGATACATGTCATATCCGTTTCTTAAGGTC 60

Qy 61 CTTTCTTGCCCTTAATGATGATGCGGGTTTACAAAACTTCCAGGGTGATGAT 120
 Db 61 CTTTCTTGCCCTTAATGATGATGCGGGTTTACAAAACTTCCAGGGTGATGAT 120

Qy 121 CTCCATGTCACCTTCCCACTCGGGTTCACATTTCTGATGTCGGTGGTCCAT 180
 Db 121 CTCCATGTCACCTTCCCACTCGGGTTCACATTTCTGATGTCGGTGGTCCAT 180

Qy 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCCATCAAGGGCTTTGGATGCCCA 240
 Db 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCCATCAAGGGCTTTGGATGCCCA 240

Qy 241 CGAGAGCTATCGGGCTGGTGATCGAGGGGATATATGTCGCCCAATGTCACCTATA 300
 Db 241 CGAGAGCTATCGGGCTGGTGATCGAGGGGATATATGTCGCCCAATGTCACCTATA 300

Qy 301 TTAATTATCTTGAATATTAATTTTGGAAAAATTAACAACCTTATCTTTGTGTA 360
 Db 301 TTAATTATCTTGAATATTAATTTTGGAAAAATTAACAACCTTATCTTTGTGTA 360

Qy 361 GGGCTTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGACACCATGTCATGTG 420
 Db 361 GGGCTTCAGCATAGATTTTCGCTTAGGGCCAGAAATGCGAGACACCATGTCATGTG 420

Qy 421 TCCACTATTTGGGACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTAATCCACT 480
 Db 421 TCCACTATTTGGGACTACCCAGAACAGATTTAAAAAATAACCAAGTAACTAATCCACT 480

Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTAATAACCGATCCTCTTAATAAA 540
 Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTAATAACCGATCCTCTTAATAAA 540

Qy 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAA 600
 Db 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAA 600

Qy 601 TTAATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTGTTATTTGTT 660
 Db 601 TTAATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTGTTATTTGTT 660

Qy 661 ATTGTTTTTATATACATTTCTTCTCTTACAAATGATGATTTTCTTCGATTTTATAAA 720
 Db 661 ATTGTTTTTATATACATTTCTTCTCTTACAAATGATGATTTTCTTCGATTTTATAAA 720

Qy 721 ATGACTATAAGTCATTTTATATTAAGAGACGCGATGCTAGATTTCTGTTCAAAATC 780
 Db 721 ATGACTATAAGTCATTTTATATTAAGAGACGCGATGCTAGATTTCTGTTCAAAATC 780

Qy 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840
 Db 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTT 840

Qy 841 TTCAAAAAAATTAAGTTATTTTCTCTTTAATAAAATGAAAAACCTTAGAAAAATAGAT 900
 Db 841 TTCAAAAAAATTAAGTTATTTTCTCTTTAATAAAATGAAAAACCTTAGAAAAATAGAT 900

Qy 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGATTAATTTATTTG 960
 Db 901 TGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGATTAATTTATTTG 960

Qy 961 GCCAGCCCCATTAATTTTAAACGAAACTGAAATGAGCGAAACCAAAATCTGAGCTAT 1020
 Db 961 GCCAGCCCCATTAATTTTAAACGAAACTGAAATGAGCGAAACCAAAATCTGAGCTAT 1020

Qy 1021 TTCTCTAGATTATGAAAAAGAGAGAGAGAAATCATGTTTAATCATGTGCC 1080

Db 1021 TTCTCTAGATTATGAAAAAGAGAGAGAGAAATCATGTTTAATCATGTGCC 1080

Qy 1081 TGAGATGTCGGTTTGGCAACGATAGCAACCGTAATCATATGATAGTGCCTACGTCA 1140
 Db 1081 TGAGATGTCGGTTTGGCAACGATAGCAACCGTAATCATATGATAGTGCCTACGTCA 1140

Qy 1141 GGTTCGGACGCTCTCGTGTATCTCACATGGCATTAATCATGCTTTGTTCAACGTTGTC 1200
 Db 1141 GGTTCGGACGCTCTCGTGTATCTCACATGGCATTAATCATGCTTTGTTCAACGTTGTC 1200

Qy 1201 TTGTTCCATGTCGCAAGCCCTGCTATTTGAAACCAAGAGATACCTACCCAAACAAT 1260
 Db 1201 TTGTTCCATGTCGCAAGCCCTGCTATTTGAAACCAAGAGATACCTACCTCCAAACAAT 1260

Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACACCGACATATGTTTCTGTAAC 1311
 Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACCGACATATGTTTCTGTAAC 1311

RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 AC AAH76340;
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. may's Ms45 promoter fragment.
 XX
 KM Ms45; male tissue; regulatory region; transcription; male fertility;
 KM hybrid seed; promoter; de.
 XX
 OS Zea mays.
 PN NC0200160997-A2.
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001MO-US004527.
 PR 15-FEB-2000; 2000US-00504487.
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 17.2%; Score 240; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 9.4e-105;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 CGTGTATCTCATGAGCATGATGCTGTTCAACGTTGATGCTGATGTCATGTC 1214
 Db 15 CGTGTATCTCATGAGCATGATGCTGTTCAACGTTGATGCTGATGTCATGTC 74


```
XX WO200160997-A2.
PN
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL,
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
XX
XX Query Match 2.9%; Score 40; DB 5; Length 40;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-08;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278
DB 1 AGGATACCTACTCTCCCAACATCATCTTACTCATGCAAC 40
XX
RESULT 9
AAH76335
ID AAH76335 standard; DNA; 30 BP.
XX
XX AAH76335;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX
XX WO200160997-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX
XX 15-FEB-2000; 2000US-00504487.
PR
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI
XX
XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
```

```
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
PS
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -152 to -181 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX
XX nucleotide sequence
XX
SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;
XX
XX Query Match 2.2%; Score 30; DB 5; Length 30;
XX Best Local Similarity 100.0%; Pred. No. 0.001;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1179 CATGCTGTTCACCGTTCGTTGTTCCA 1208
DB 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
XX
XX ABL40437/c
ID ABL40437 standard; DNA; 849 BP.
XX
XX ABL40437;
AC
XX
XX 10-JUN-2002 (first entry)
DT
XX
XX Sunflower chitinase promoter sequence.
DE
XX
XX Sunflower; chitinase; lipid transfer protein; LTP; transcription;
KM sclerotinia; genetic engineering; disease resistance; insecticide;
KM antifungal; viricide; nematocide; antimicrobial; antibacterial;
KM gene therapy; gene; promoter; ds.
XX
XX Helianthus annuus.
OS
XX
XX Key Location/Qualifiers
FH CHAT_signal 723..726
FT /*tag= a
FT TATA_signal 807..811
FT /*tag= b
XX
XX WO200214502-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 08-AUG-2001; 2001WO-US041629.
PF
XX
XX 11-AUG-2000; 2000US-0224603P.
PR
XX
XX 07-AUG-2001; 2001US-00923844.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Bao Z, Lu G;
PI
XX
XX WPI; 2002-269194/31.
DR
XX
XX New genes and promoters from Helianthus annuus, for genetically
PT manipulating plants to enhance disease resistance or resistance to
PT pathogen, and for producing proteins for controlling plant diseases
PT caused by these pathogens.
XX
XX Claim 1; Fig 4; 89pp; English.
PS
XX
XX The invention relates to sclerotinia-inducible genes and promoters capable
CC of initiating transcription from sunflower. The polynucleotides are
```


CC useful for genetically manipulating plants to enhance disease resistance,
CC or resistance to insect, fungal pathogens, viruses or nematodes. The
CC nucleotide sequences are also useful as genetic markers in disease-
CC resistance breeding programs. The polynucleotides are also useful for
CC producing or eliminating useful in compositions. The compositions are useful
CC reducing or eliminating damage to plants caused by plant pathogens. The
CC compositions find use in agricultural and pharmaceutical compositions as
CC antifungal and antimicrobial agent. For agricultural purposes, the
CC compositions may be used in sprays for control of plant disease. As
CC pharmaceutical compositions, the agents are useful as antibacterial and
CC antimicrobial treatments. The present sequence represents the promoter
CC fragment of the sunflower chitinase gene
XX

Sequence 849 BP; 295 A; 139 C; 135 G; 280 T; 0 U; 0 Other;

Query Match 1.6%; Score 23; DB 6; Length 849;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 AGTTATTTCTCTTATAAAT 876
|||||
Db 556 AGTTATTTCTCTTATAAAT 534

RESULT 11
ABV56505
ID ABV56505 standard; cDNA; 333 BP.
XX
XX ABV56505;

AC
XX
XX 17-SEP-2002 (first entry)
XX
XX

DE Human prostate expression marker cDNA 56496.
XX

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.

OS Homo sapiens.
XX

PN WO200160860-A2.
XX

PD 23-AUG-2001.
XX

PF 20-FEB-2001; 2001WO-US005171.
XX

PR 17-FEB-2000; 2000US-0183119P.
XX

PR 16-MAR-2000; 2000US-0189862P.
XX

PR 25-MAY-2000; 2000US-0207454P.
XX

PR 09-JUN-2000; 2000US-0211314P.
XX

PR 18-JUL-2000; 2000US-0219007P.
XX

PR 13-DEC-2000; 2000US-0255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX

PI Schlegel R, Endege WO, Monahan JE;
XX

DR WPI; 2001-662795/76.
XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 10900; 11750pp; English.
XX

CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

Sequence 333 BP; 75 A; 69 C; 46 G; 143 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 5; Length 333;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 TTACAAAAAAGCTTCACGGGT 113
|||||
Db 89 TTACAAAAAAGCTTCACGGGT 109

RESULT 12
ABA58867/C
ID ABA58867 standard; DNA; 459 BP.
XX

AC ABA58867;
XX

DT 01-FEB-2002 (first entry)
XX

DE Human foetal liver single exon nucleic acid probe #7172.
XX

KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX

OS Homo sapiens.
XX

PN WO200157277-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US000669.
XX

PR 04-FEB-2000; 2000US-0180312P.
XX

PR 26-MAY-2000; 2000US-0207456P.
XX

PR 30-JUN-2000; 2000US-00608408.
XX

PR 03-AUG-2000; 2000US-00632366.
XX

PR 21-SEP-2000; 2000US-0234687P.
XX

PR 27-SEP-2000; 2000US-0236359P.
XX

PR 04-OCT-2000; 2000GB-00024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-483447/52.
XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.

PS Claim 1; SEQ ID NO 7172; 639pp + Sequence Listing; English.
XX

CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX

Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
|||||
Db 118 AAAGGAGAGAGAGAGAGA 98


```
RESULT 13
AA138581/c
ID AA138581 standard; DNA; 459 BP.
XX
AC AA138581;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #7267 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 7267; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98

RESULT 14
AAK32768/c
ID AAK32768 standard; DNA; 459 BP.
XX
AC AAK32768;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 7325.
XX
KM Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
```

```
PN WO200157276-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000668.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-00632366.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX DR WPI; 2001-488900/53.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
XX PS Example 4; SEQ ID NO 7325; 658bp + Sequence Listing; English.
XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
AAK07027/c
ID AAK07027 standard; DNA; 459 BP.
XX
AC AAK07027;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 7018.
XX
XX KM Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KM Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX KM ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000667.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-00632366.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
```


PA (MOL-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 DR
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 XX
 PS Example 4; SEQ ID NO 7018; 650pp + Sequence listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention
 XX
 SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
 |||||
 Db 118 AAAGGAGAGAGAGAGAGA 98

Search completed: September 16, 2005, 13:08:40
 Job time : 579.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 19:42:02 ; Search time 176.5 Seconds
(without alignments)
12923.341 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394

Sequence: 1 cccctggtcgtctcatgaaga.....cttttcgtccaccaccatg 1394

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: /cgcn2_6/pdata/1/lna/5A.COMB.seq.*
2: /cgcn2_6/pdata/1/lna/5B.COMB.seq.*
3: /cgcn2_6/pdata/1/lna/6A.COMB.seq.*
4: /cgcn2_6/pdata/1/lna/6B.COMB.seq.*
5: /cgcn2_6/pdata/1/lna/6C.COMB.seq.*
6: /cgcn2_6/pdata/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1394	100.0	1394	US-08-880-499-2	Sequence 2, Appl1
2	1311	94.0	1394	US-08-880-499-1	Sequence 1, Appl1
3	21	1.5	50263	US-09-949-016-13563	Sequence 13563, A
4	20	1.4	601	US-09-949-016-12481	Sequence 12481, A
5	20	1.4	2523	US-08-410-784A-3	Sequence 14705, A
6	20	1.4	21679	US-09-949-016-15250	Sequence 15250, A
7	20	1.4	65424	US-09-949-016-12426	Sequence 12426, A
8	20	1.4	374159	US-09-949-016-15868	Sequence 15868, A
9	20	1.4	601	US-09-949-016-17933	Sequence 17933, A
10	19	1.4	601	US-09-949-016-11823	Sequence 41823, A
11	19	1.4	601	US-09-949-016-11824	Sequence 14534, A
12	19	1.4	601	US-09-949-016-145344	Sequence 145345, A
13	19	1.4	601	US-09-949-016-145345	Sequence 2409, Ap
14	19	1.4	768	US-08-328-352-2409	Sequence 10, Appl
15	19	1.4	1584	US-08-247-908A-10	Sequence 10, Appl
16	19	1.4	1584	US-08-453-942-10	Sequence 10, Appl
17	19	1.4	1584	US-08-926-885A-10	Sequence 10, Appl
18	19	1.4	1584	PCT-US94-05290-10	Sequence 10, Appl
19	19	1.4	1833	US-08-365-470-2	Sequence 2, Appl1
20	19	1.4	1833	US-08-209-668-18	Sequence 18, Appl
21	19	1.4	3834	US-09-949-016-71	Sequence 71, Appl
22	19	1.4	3834	US-08-365-470-1	Sequence 1, Appl1
23	19	1.4	3854	US-08-344-155C-98	Sequence 98, Appl
24	19	1.4	3858	US-09-009-490A-88	Sequence 88, Appl
25	19	1.4	3858	US-09-949-016-4401	Sequence 4401, Ap
26	19	1.4	3862		
27	19	1.4	3862		

C 28	19	1.4	3863	3	US-08-482-073-1	Sequence 1, Appl1
C 29	19	1.4	3863	6	5217870-1	Patent No. 5217870
C 30	19	1.4	3863	6	5217870-1	Patent No. 5217870
C 31	19	1.4	9370	1	US-08-320-559-27	Sequence 27, Appl1
C 32	19	1.4	9370	3	US-08-545-860D-27	Sequence 27, Appl1
C 33	19	1.4	9370	5	PCT-US94-04496-27	Sequence 27, Appl1
C 34	19	1.4	9391	1	US-08-320-559-25	Sequence 25, Appl1
C 35	19	1.4	9391	3	US-08-545-860D-25	Sequence 25, Appl1
C 36	19	1.4	9391	5	PCT-US94-04496-25	Sequence 25, Appl1
C 37	19	1.4	15384	4	US-09-949-016-11813	Sequence 11813, A
C 38	19	1.4	15413	4	US-09-949-016-16143	Sequence 16143, A
C 39	19	1.4	71645	4	US-09-949-016-12126	Sequence 12126, A
C 40	19	1.4	71651	4	US-09-949-016-17258	Sequence 17258, A
C 41	19	1.4	105189	4	US-09-949-016-13029	Sequence 13029, A
C 42	19	1.4	133559	4	US-09-949-016-15845	Sequence 15845, A
C 43	19	1.4	144158	4	US-09-949-016-11755	Sequence 11755, A
C 44	19	1.4	144158	4	US-09-949-016-12936	Sequence 12936, A
C 45	19	1.4	228896	4	US-09-949-016-17127	Sequence 17127, A

ALIGNMENTS

RESULT 1
US-08-880-499-2
Sequence 2, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Releasee #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-2

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAAGGCTC 60
DB 1 CCATGCTCTCTATGAAAAAGATGATACATGTCCTATATCCGTTTCTTAAAGGCTC 60
QY 61 CTTCTTCTGCTTATTACTGATGATCGGGGTTTACAAAACTTCCACGGGTGCATGAT 120
DB 61 CTTCTTCTGCTTATTACTGATGATCGGGGTTTACAAAACTTCCACGGGTGCATGAT 120
QY 121 CTCATGTTCCACTTCTCCACCTCGGCTGACATTTCTTGATGTCGGTGGTCCCAT 180
DB 121 CTCATGTTCCACTTCTCCACCTCGGCTGACATTTCTTGATGTCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGCGCA 240
DB 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCATCAAGGACCTTTCGATGCGCA 240
QY 241 CGAGAGCTATCGGGTCGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
DB 241 CGAGAGCTATCGGGTCGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
QY 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGATGTA 360
DB 301 TTATTTATCTTTAGATATTTATTTTGAAGAAATACAACTTATCTTTGATGTA 360
QY 361 GGGCCTCAGCATGATTTTCGCTTAAAGGCCAGAAATGCGAGAGCAGCCATGCTATG 420
DB 361 GGGCCTCAGCATGATTTTCGCTTAAAGGCCAGAAATGCGAGAGCAGCCATGCTATG 420
QY 421 TCCACTTATGCGACTACCCGAAACAGATTTAAAAAATACCAAGTAACTATCACT 480
DB 421 TCCACTTATGCGACTACCCGAAACAGATTTAAAAAATACCAAGTAACTATCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCCTTAAAAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCCTTAAAAA 540
QY 541 CAAGCATATTTGCAAGAGACAAATATGTTACAGTTTACAAACATCTAAGCGACAA 600
DB 541 CAAGCATATTTGCAAGAGACAAATATGTTACAGTTTACAAACATCTAAGCGACAA 600
QY 601 TTATATCGAAGAGTAACTATGACGTTCAAGATTTTCTTTTCACTTCTGTTATTTGTT 660
DB 601 TTATATCGAAGAGTAACTATGACGTTCAAGATTTTCTTTTCACTTCTGTTATTTGTT 660
QY 661 ATTGTTTATATACATTTTCTTCTTATACATAGATGTTTCTTCCGATTTTATAA 720
DB 661 ATTGTTTATATACATTTTCTTCTTATACATAGATGTTTCTTCCGATTTTATAA 720
QY 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGCTAGATTTCTGTTCAAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATAAAGACGACATGCTAGATTTCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTAAAGAGTAACTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCTGATTTTAAAGAGTAACTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAACTTGAAGATGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAACTTGAAGATGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTAATATTTTGG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTAATATTTTGG 960
QY 961 GCGAGCCCATTAATTTTAAACCGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
DB 961 GCGAGCCCATTAATTTTAAACCGAACTGAATCGAGCGAAACCAATCTGACTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
QY 1081 TGAGATGTGGGTTTGGCAACATAGCCAGTAAATCATAGTCAATAGTGTGCTACGTCA 1140

DB 1081 TGAGATGTGGGTTTGGCAACATAGCCAGCTTAATACAGTCAATAGTGTGCTACGTCA 1140
QY 1141 GGTTCGGACGCTCTGCTGATCTCATGCAATGCAATACATGCTTTTCAACCGTTGCTC 1200
DB 1141 GGTTCGGACGCTCTGCTGATCTCATGCAATGCAATACATGCTTTTCAACCGTTGCTC 1200
QY 1201 TTGTTTCATGCTCAAGCCCTGCTATTCTGAACCAAGAGATPACTACTCCCAACAT 1260
DB 1201 TTGTTTCATGCTCAAGCCCTGCTATTCTGAACCAAGAGATPACTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTTCATGCAACACGCAATATGTTTCTGAAACAGATCTATT 1320
DB 1261 CCATCTTACTCATGCAACTTTCATGCAACACGCAATATGTTTCTGAAACAGATCTATT 1320
QY 1321 AAAAGTCAACAAAGCTAGGTTCTCCCGTACTTCCCTCTCTCTGCGCATCTTTT 1380
DB 1321 AAAAGTCAACAAAGCTAGGTTCTCCCGTACTTCCCTCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACGACGATG 1394
DB 1381 CGTCCACGACGATG 1394

RESULT 2
US-08-880-499-1
Sequence 1, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garnat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
NUMBER OF SEQUENCES: 2
TITLE OF INVENTION: AND METHOD OF USING SAME
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880 499
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-1
Query Match 94.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Qy 1 CCATGCTCTCTATGAAAAAGATGATACAACTGCTCTATATCCGTTTCTTAGGCTCC 60
Db 1 CCATGCTCTCTATGAAAAAGATGATACAACTGCTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCCTTCGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 61 CTTCCTTCGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 121 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Db 121 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 181 CTGACGAGGCGCATGACGACCTTTGGGACCCCAATGAGGCGCTTGGATGGCCCA 240
Db 181 CTGACGAGGCGCATGACGACCTTTGGGACCCCAATGAGGCGCTTGGATGGCCCA 240
Qy 241 CGAGACGCTATCGGCTCGTGATCCAGGGATATATGTCCTCCCAATGCTCACTATA 300
Db 241 CGAGACGCTATCGGCTCGTGATCCAGGGATATATGTCCTCCCAATGCTCACTATA 300
Qy 301 TTATATCTTTAGATATTAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Db 301 TTATATCTTTAGATATTAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Qy 361 GGGCTTCAGATGATTTGGCTTAGGGCCCAAGAAAGGACGACGATGCTTAGTG 420
Db 361 GGGCTTCAGATGATTTGGCTTAGGGCCCAAGAAAGGACGACGATGCTTAGTG 420
Qy 421 TCCATATTTGGCATACCCAGAACAGATTTTAAATAAATTAACAACTTATCTTAGT 480
Db 421 TCCATATTTGGCATACCCAGAACAGATTTTAAATAAATTAACAACTTATCTTAGT 480
Qy 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTAACCAATCTCTTAAAAA 540
Db 481 CGAAGCTATCATGATATGTTTAAAGAAACATCTATTAACCAATCTCTTAAAAA 540
Qy 541 CAAGCATATTTCCAGAAAGACAAATTTATGTTACAGTTTAAACATCTTAGACGACAA 600
Db 541 CAAGCATATTTCCAGAAAGACAAATTTATGTTACAGTTTAAACATCTTAGACGACAA 600
Qy 601 TTATATCGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 TTATATCGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 661 ATTGTTTTATATACATTTTCTCTTAAATAGAGTATTTTCCGATTTTAAAA 720
Db 661 ATTGTTTTATATACATTTTCTCTTAAATAGAGTATTTTCCGATTTTAAAA 720
Qy 721 ATGACTATAAGCTATTTTATATATAGAGACGATGCTGATGATTTCTGAAAAATC 780
Db 721 ATGACTATAAGCTATTTTATATATAGAGACGATGCTGATGATTTCTGAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTCAAAAAAATAGTTATTTTCTCTTAAATAGAAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTATTTTCTCTTAAATAGAAACACTTAGAAAAATAGAGT 900
Qy 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTGATATTTATTTG 960
Db 901 TGGCAGCTAGCCCTAGAAATGTTTCCCAATTAATTCATCTGCTGATATTTATTTG 960
Qy 961 GCCAGCCCATTAATTTTAAACCGAAATGAAATGAGCGGAAACCAATCTGAGTAT 1020
Db 961 GCCAGCCCATTAATTTTAAACCGAAATGAAATGAGCGGAAACCAATCTGAGTAT 1020
Qy 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

```

```

Qy 1081 TGAGATGTGGGCTTTGGCAACGATAGCACCCGTAATCATAGCTATAGGTCCTAGCTCA 1140
Db 1081 TGAGATGTGGGCTTTGGCAACGATAGCACCCGTAATCATAGCTATAGGTCCTAGCTCA 1140
Qy 1141 GGTTCGGACGCTCTCGTGTATCTCATATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1141 GGTTCGGACGCTCTCGTGTATCTCATATGATGATGATGATGATGATGATGATGATGAT 1200
Qy 1201 TTGTTCCATGCTGCAAGCCTTGCTTATTTGAAACGAAGATGATGATGATGATGATGAT 1260
Db 1201 TTGTTCCATGCTGCAAGCCTTGCTTATTTGAAACGAAGATGATGATGATGATGATGAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACGACGACATATGTTCTGTAAC 1311
Db 1261 CCATCTTACTCATGCACTTCCATGCAACGACGACATATGTTCTGTAAC 1311

RESULT 3
US-09-949-016-13563
; Sequence 13563: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13563
; LENGTH: 50263
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13563

Query Match
Best local Similarity 100.0%; Score 21; DB 4; Length 50263;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 861 TTCTCTTATTAATAATAGAAA 881
Db 40441 TTCTCTTATTAATAATAGAAA 40461

RESULT 4
US-09-949-016-124481/c
; Sequence 124481: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124481
; LENGTH: 601
; TYPE: DNA

```


ORGANISM: Human
US-09-949-016-124481

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
DB 363 GATTTTCTTTTCATCTT 344

RESULT 5

US-09-949-016-147075/c
Sequence 147075, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 147075
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-147075

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 TTTCCTTCAAGATTTCGA 836
DB 353 TTTCCTTCAAGATTTCGA 334

RESULT 6

US-08-410-784A-3/c
Sequence 3, Application US/08410784A
Patent No. 5912413
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Ph.D., Holiday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:

US-08-410-784A-3

Query Match 1.4%; Score 20; DB 2; Length 2523;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 849 AATTAGTTTATTTCTCTT 868
DB 206 AATTAGTTTATTTCTCTT 187

RESULT 7

US-09-949-016-15250
Sequence 15250, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15250
LENGTH: 21679
TYPE: DNA
ORGANISM: Human
US-09-949-016-15250

Query Match 1.4%; Score 20; DB 4; Length 21679;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCATCTT 649
DB 4880 GATTTTCTTTTCATCTT 4899

RESULT 8

US-09-949-016-12426/c
Sequence 12426, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.


```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12426
; LENGTH: 65424
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(65424)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12426

Query Match          1.4%; Score 20; DB 4; Length 65424;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1041 GGAGAGAGAGAGAGAGAAAT 1060
Db      50936 GGAGAGAGAGAGAGAGAAAT 50917

RESULT 9
US-09-949-016-15868
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Query Match          1.4%; Score 20; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      817 TTCTCTTCAAAAGATTGCA 836
Db      265055 TTCTCTTCAAAAGATTGCA 265074

RESULT 10
US-09-949-016-17933/C
; Sequence 17933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17933

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAAAAATTAACCAAG 467
Db      361 TTTAAAAAATTAACCAAG 343

RESULT 11
US-09-949-016-17934/C
; Sequence 17934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17934

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAAAAATTAACCAAG 467
Db      394 TTTAAAAAATTAACCAAG 376

RESULT 12
US-09-949-016-41823/C
; Sequence 41823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```



```
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41823
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-41823

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATACCAAG 467
DB      361 TTTAAATAATACCAAG 343

RESULT 13
US-09-949-016-41824/c
/ Sequence 41824, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR FILING DATE: 2000-09-08
/ PRIOR APPLICATION NUMBER: 60/231,498
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41824
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-41824

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      449 TTTAAATAATACCAAG 467
DB      394 TTTAAATAATACCAAG 376

RESULT 14
US-09-949-016-145344/c
/ Sequence 145344, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
```

```
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 145344
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-145344

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAG 1056
DB      561 AAGGAGAGAGAGAGAG 543

RESULT 15
US-09-949-016-145345/c
/ Sequence 145345, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 145345
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-145345

Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1038 AAGGAGAGAGAGAGAG 1056
DB      553 AAGGAGAGAGAGAGAG 535
```

Search completed: September 16, 2005, 08:14:17
Job time : 178.5 secs


```
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTGTCTATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCTCACGGGTGCATGAT 120
Db 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCTCACGGGTGCATGAT 120
Qy 121 CTCGATGTTTCACCTTCTCCCACTCGGGTTCACATTTCTTGGATGCGTGGTCCCAT 180
Db 121 CTCGATGTTTCACCTTCTCCCACTCGGGTTCACATTTCTTGGATGCGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCATCAAGGGCCTTTGCAATGCGCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCATCAAGGGCCTTTGCAATGCGCA 240
Qy 241 CGAGAGGTATCGGGTGTGTGTATCCAGGGATATATGTCCCAAAATCGTCACTTAA 300
Db 241 CGAGAGGTATCGGGTGTGTGTATCCAGGGATATATGTCCCAAAATCGTCACTTAA 300
Qy 301 TTATTAATCTTATGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTAATCTTATGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGCTCTACAGATGATTTTGGCTTAGGGCCAGAAATGCGAGGACCAAGCCATGCTAGTG 420
Db 361 GGGCTCTACAGATGATTTTGGCTTAGGGCCAGAAATGCGAGGACCAAGCCATGCTAGTG 420
Qy 421 TCCGATTTGGGCACTCCAGAACAAAGTTTAAATTAACCAAGTAACTTAATCCACT 480
Db 421 TCCGATTTGGGCACTCCAGAACAAAGTTTAAATTAACCAAGTAACTTAATCCACT 480
Qy 481 CGAAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACCAAGCATCCTTAAAAA 540
Db 481 CGAAAGCTATCATGATATGTTTAAAGAAACATCTATTTAAACCAAGCATCCTTAAAAA 540
Qy 541 CAAAGCATTTTGAAGAGACAAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
Db 541 CAAAGCATTTTGAAGAGACAAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
Qy 601 TTAATATCGAAAGGTAGCTATGAGCTTCAAGATTTTCTTTTCAATCTTATTTTGT 660
Db 601 TTAATATCGAAAGGTAGCTATGAGCTTCAAGATTTTCTTTTCAATCTTATTTTGT 660
Qy 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGATTTTCTCCGATTTTATAA 720
Db 661 ATTGTTTTATATACATTTTCTTCTTCAATAGAGATTTTCTCCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATTAAGACACGATGCTGATGATCTCGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATTAAGACACGATGCTGATGATCTCGTTCAAAAATC 780
Qy 781 TTTTGAATTTTTTAAGAGCTAGTTGGCAACCCCTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTTGAATTTTTTAAGAGCTAGTTGGCAACCCCTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATTAAGTTTATTTTCTTATTAATAATAGAAAACCTTAGAAAAATAGAT 900
Db 841 TTTCAAAAAAATTAAGTTTATTTTCTTATTAATAATAGAAAACCTTAGAAAAATAGAT 900
Qy 901 TGGCAGACTAGCCCTGAAATGTTTTCCCAATAAATTAACAATCACTGTGATTAATTTTG 960
Db 901 TGGCAGACTAGCCCTGAAATGTTTTCCCAATAAATTAACAATCACTGTGATTAATTTTG 960
Qy 961 GCGAGCCCATTAATTAATTTAAACGAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTAATTTAAACGAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGCCTTACGTCA 1140
Db 1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGCCTTACGTCA 1140
```

```
Db 1081 TGAAGATGCGGTTTGGCAACGATAGCCACCGTAATCATAGTCAATAGTGCCTTACGTCA 1140
Qy 1141 GGTTCGGGAGCTCTGTGTGATCTCATAGAGATATACATCTTGTTCACCGTTCGTC 1200
Db 1141 GGTTCGGGAGCTCTGTGTGATCTCATAGAGATATACATCTTGTTCACCGTTCGTC 1200
Qy 1201 TTGTTTCATGCTCAGAGCTTCTCTATTTCTGAACCAAGAGATATCTACTCCCAACAAT 1260
Db 1201 TTGTTTCATGCTCAGAGCTTCTCTATTTCTGAACCAAGAGATATCTACTCCCAACAAT 1260
Qy 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAAACAGATCTATT 1320
Db 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAAACAGATCTATT 1320
Qy 1321 AAAGATCAACAAGTAGGTTCTCCGCTACTTCCCTCTCCCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGTAGGTTCTCCGCTACTTCCCTCTCCCTGCGCATCTTTT 1380
Qy 1381 CGTCCACGACCATG 1394
Db 1381 CGTCCACGACCATG 1394
```

```
RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713.381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1
```

```
Query Match 94.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTGTCTATCCGTTTCTTAGGGTCC 60
Db 1 CCAATGCTCTATGAAAAAGATGAGTCAATGTGTCTATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCTCAAGGGTGCATGAT 120
Db 61 CTTCTCTGCTTATTAAGTCACTGAATCGGGGTTACAAAAAATTCTCAAGGGTGCATGAT 120
Qy 121 CTCGATGTTTCACCTTCTCCCACTCGGGTTCACATTTCTTGGATGCGTGGTCCCAT 180
Db 121 CTCGATGTTTCACCTTCTCCCACTCGGGTTCACATTTCTTGGATGCGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCATCAAGGGCCTTTGCAATGCGCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTGCGGACACCATCAAGGGCCTTTGCAATGCGCA 240
Qy 241 CGAGAGGTATCGGGTGTGTGTATCCAGGGATATATGTCCCAAAATCGTCACTTAA 300
Db 241 CGAGAGGTATCGGGTGTGTGTATCCAGGGATATATGTCCCAAAATCGTCACTTAA 300
Qy 301 TTATTAATCTTATGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTAATCTTATGATATTAATTTTGGAAAAATACAACTTATCTTTGTGTA 360
```



```

Db 301 TTATTAATCTTTAGATATTTAATTTTGGAAAAATACAACTTATCTTTGGTGA 360
Qy 361 GGGCCCTCAGATATGATTTTCGTTAGGGCCAGAAATCGAGAGACGAGCCATCTAGTG 420
Db 361 GGGCCCTCAGATATGATTTTCGTTAGGGCCAGAAATCGAGAGACGAGCCATCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480
Db 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAA 540
Qy 541 CAAGCATATTTCCAAAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAAGCGACAA 600
Db 541 CAAGCATATTTCCAAAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAAGCGACAA 600
Qy 601 TTATATCGAAAGGTAAGCTATGAGTTCAGATTTTCTTTCAATCTGTTATTTTGT 660
Db 601 TTATATCGAAAGGTAAGCTATGAGTTCAGATTTTCTTTCAATCTGTTATTTTGT 660
Qy 661 ATTTGTTTATATACATTTTCTTCTCAATAGAGTATTTCTTCGATTTTATTA 720
Db 661 ATTTGTTTATATACATTTTCTTCTCAATAGAGTATTTCTTCGATTTTATTA 720
Qy 721 ATGACTATTAAGCATTTTATTAATTAAGAGACGAGATGCTAGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGCATTTTATTAATTAAGAGACGAGATGCTAGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTTGGCAACGCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACGCTGTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAATAGAGT 900
Db 841 TTTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATCTAGAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTGAGTAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTTTG 960
Db 901 TGGCAGACTAGCCCTGAGTAATGTTTCCCAATAAATTAACAATCACTGTGTATTAATTTTG 960
Qy 961 GCGAGCCCATTAATTTAATTAACCGAAATCGAAATCGAGCAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTTAATTAACCGAAATCGAAATCGAGCAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTC 1080
Db 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTC 1080
Qy 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTCACTGA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTCACTGA 1140
Qy 1141 GGTTCGGGAGCTCTCGTGCATCTTCAATGAGCTACTAATGTTTCAACCGTTCGTC 1200
Db 1141 GGTTCGGGAGCTCTCGTGCATCTTCAATGAGCTACTAATGTTTCAACCGTTCGTC 1200
Qy 1201 TTGTTCCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATCTTATCCCAACAT 1260
Db 1201 TTGTTCCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATCTTATCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311

```

RESULT 3
 US-10-713-381-9
 ; Sequence 9, Application US/10713381
 ; Publication No. US20040221331A1
 ; GENERAL INFORMATION:

```

; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 17.2%; Score 240; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 5.8e-110;
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 CGTGCATCTCATGSCATCTACTACATGCTTTGCAACGCTTCGTTGTCATGCTCC 1214
Db 15 CGTGCATCTCATGSCATCTACTACATGCTTTGCAACGCTTCGTTGTCATGCTCC 74

Qy 1215 AAGCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAAATCCATCTTACTCATG 1274
Db 75 AAGCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAAATCCATCTTACTCATG 134

Qy 1275 CAATTCATGCAAAACGCGACATATGTTTCTGTAACAGATTAATTAAGATCAACAAG 1334
Db 135 CAATTCATGCAAAACGCGACATATGTTTCTGTAACAGATTAATTAAGATCAACAAG 194

Qy 1335 CTAGGCTTCCCGCTAGGTTCCCTCTCCCTGCGCATCTTTTTCGTCACACCATG 1394
Db 195 CTAGGCTTCCCGCTAGGTTCCCTCTCCCTGCGCATCTTTTTCGTCACACCATG 254

```

RESULT 4
 US-10-713-381-3
 ; Sequence 3, Application US/10713381
 ; Publication No. US20040221331A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, MARC C.
 ; APPLICANT: FOX, TIMOTHY W.
 ; APPLICANT: GARNAT, CARL W.
 ; APPLICANT: HUFFMAN, GARY
 ; APPLICANT: KENDALL, TIMMY L.
 ; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
 ; FILE REFERENCE: 578R
 ; CURRENT APPLICATION NUMBER: US/10/713,381
 ; PRIOR FILING DATE: 2003-11-14
 ; PRIOR APPLICATION NUMBER: 08/880,499
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 158
 ; TYPE: DNA
 ; ORGANISM: Zea mays

US-10-713-381-3
 Query Match 8.0%; Score 111; DB 20; Length 158;
 Best Local Similarity 100.0%; Pred. No. 6.2e-45;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1201 TTGTTCCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATCTTATCCCAACAT 1260
Db 48 TTGTTCCATCTGTCAGAGCTTGGCTATTTCTGAACCAAGAGATCTTATCCCAACAT 107

```


Oy 1261 CCATCTTACTGATGCAACTTCATGCAACAACGCACATATGTTCCGTGAAC 1311
|||||
Db 108 CCACTTACTCATGCAACTTCCATGCCAACAACGCACATATGTTCCGTGAAC 158

```

RESULT 5
US-10-739-930-1857
; Sequence 1857, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377) B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1857
; LENGTH: 1663
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

```

QY 1318 ATTAAAGATCACAAACAGCTAGCGTTCTCCCGCTAGCTTCCCTCTCTCTCTGCGATCTT 1377

Db 6 ATTAAAGATCACAAACAGCTAGCGTTCTCCCGCTAGCTTCCCTCTCTCTCTGCGATCTT 65

QY	1378	TTTTCG	1382
Db	66	TTTTCG	70

```

RESULT 6
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARRAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

```

Query Match	3.6%	Score 50;	DB 20;	Length 50;
Best Local Similarity	100.0%	Pred. No. 3.3e-14;		
Matches 50; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

[illegible]

```

RESULT 7
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331a1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARMAAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 576R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

```

QY	1239	AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC	1278
Db	1	AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC	40

```

RESULT 8
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARMAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713.381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

```

Query Match	2.2%	Score 30;	DB 20;	Length 30;
Best Local Similarity	100.0%	Pred. No. 0.00038;		
Matches 30; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1179	CATGCTGTCAACCGTTCGCTTGTCCA	1208
Db	1	CATGCTGTCAACCGTTCGCTTGTCCA	30

RESULT 9
US-09-923-844B-5/c
; Sequence 5, Application US/09923844B
; Patent No. US20020166143A1
; GENERAL INFORMATION:


```
; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Bao, Zhongmeng
; APPLICANT: Lu, Guhua
; TITLE OF INVENTION: Sclerotinia-inducible Genes and
; FILE REFERENCE: 35718/234631
; CURRENT APPLICATION NUMBER: US/09/923,844B
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/224,603
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Helianthus annuus
; US-09-923-844B-5

Query Match      1.6%; Score 23; DB 9; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      854 AGTTATTTCTCTTTATAAAT 876
Db      556 AGTTATTTCTCTTTATAAAT 534

RESULT 10
US-10-085-117-355/C
; Sequence 355, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 355
; LENGTH: 194945
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(194945)
; OTHER INFORMATION: n = any nucleotide
; US-10-085-117-355

Query Match      1.6%; Score 22; DB 17; Length 194945;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 TGCATGATCTCATGTTTCACT 134
Db      28972 TGCATGATCTCATGTTTCACT 28951

RESULT 11
US-10-357-930-56524
; Sequence 56524, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
```

```
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 6232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56524
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-357-930-56524

Query Match      1.5%; Score 21; DB 20; Length 333;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      93 TTACAAAAAAGTTCACGGGT 113
Db      89 TTACAAAAAAGTTCACGGGT 109

RESULT 12
US-10-425-115-183040
; Sequence 183040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 183040
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98517C.1
; US-10-425-115-183040

Query Match      1.5%; Score 21; DB 20; Length 356;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      632 TTTTCTTTTCATTTCTTGT 652
Db      72 TTTTCTTTTCATTTCTTGT 92

RESULT 13
US-10-425-115-117183/C
; Sequence 117183, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```



```
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 117183
/ LENGTH: 386
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_3835C.1
US-10-425-115-117183

Query Match          1.5%; Score 21; DB 20; Length 386;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      371 ATAGATTTGCTTAGGGCCC 391
DB      367 ATGATTTTGCCTTAGGGCCC 347

RESULT 14
US-09-864-761-11465/C
/ Sequence 11465, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wansheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
```

```
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Aniomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 11465
/ LENGTH: 459
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC000100.2
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
US-09-864-761-11465

Query Match          1.5%; Score 21; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
DB      118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
US-10-425-115-51470/C
/ Sequence 51470, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 51470
/ LENGTH: 1203
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(1203)
/ OTHER INFORMATION: unsure at all n locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470

Query Match          1.5%; Score 21; DB 20; Length 1203;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Search completed: September 16, 2005, 12:49:11
Job time : 1293.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 10:24:35 ; Search time 4293.5 Seconds

(without alignments)
12358.587 Million cell updates/sec

Title: US-10-713-381-2

Sequence: 1394
1 cccctggctgcctcctatgaaaa.....cttttcgccaccaccatcg 1394

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	908	65.1	963	9	CC656933 OGMDO20TM
2	667	47.8	915	9	CG224225 OGIAG08TV
3	409	29.3	687	9	CC656939 OGMDO20TV
4	66	4.7	878	9	CG224211 OGIAG08TV
5	27	1.9	539	8	B2828689 B2828689
6	27	1.9	611	8	B2828694 B2828694
7	27	1.9	804	9	CG148339 PUBG46TV
8	26	1.9	854	9	CG011564 ZUABT3TV
9	25	1.8	687	8	B2371637 B2371637
10	25	1.8	726	9	CC810403 ZMMBC047
11	25	1.8	819	9	CG305519 OCGMDO2TV
12	25	1.8	825	9	CG190606 PUBH82TV
13	25	1.8	833	9	CG190609 PUBH82TV
14	25	1.8	857	8	CC358144 PCHPY12TV
15	25	1.8	878	9	CG307825 OCMBC01TV
16	25	1.8	885	8	CG307825 OCMBC01TV
17	25	1.8	892	9	CG197610 PUBH81TV
18	25	1.8	893	8	CC439146 PUBH86TV
19	25	1.8	956	9	CG060573 PUBCN7TV
20	24	1.7	385	1	A1626382 A1626382
21	24	1.7	468	1	A1626141 A1626141
22	24	1.7	502	8	BH411902 BH411902
23	24	1.7	533	8	CC010354 CC010354
24	24	1.7	625	9	CG377165 CG377165

25	24	1.7	630	9	CG331717 CG331717
26	24	1.7	699	9	CM009494 CM009494
27	24	1.7	739	9	CM009496 CM009496
28	24	1.7	835	9	CC979335 ZUACG47TV
29	24	1.7	857	8	CC335124 OGUCH74TV
30	24	1.7	866	9	CG209882 OGIAB01TV
31	23	1.6	229	8	CC364841 PUBH06TV
32	23	1.6	318	9	CG421210
33	23	1.6	493	1	AL375883
34	23	1.6	514	8	B2032120 B2032120
35	23	1.6	533	8	AZ095114 AZ095114
36	23	1.6	610	4	B1676219 B1676219
37	23	1.6	651	8	BH297318 BH297318
38	23	1.6	664	9	CG700282
39	23	1.6	691	8	B2524893 B2524893
40	23	1.6	710	8	B2828501 B2828501
41	23	1.6	831	8	B2129958 B2129958
42	23	1.6	833	9	CG437195 CG437195
43	23	1.6	845	9	CG255896 CG255896
44	23	1.6	870	9	CL198382 CL198382
45	23	1.6	900	9	CT708844 CT708844

ALIGNMENTS

RESULT 1	CC656933	963 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDO20TM_ZM_0.7_1.5_KB	Zea mays	genomic clone	ZMMBMA0554D15,	
DEFINITION	genomic survey sequence.				
ACCESSION	CC656933	GI:32060225			
VERSION	CC656933.1				
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utecherback, T., Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 963)				
AUTHORS	Whitelaw, C.A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: OGMDO20TV Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: TR Class: sheared ends.				
FEATURES	Location/Qualifiers				
source	1..963 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMA0554D15" /clone_lib="ZM_0.7_1.5_KB" /note="Vector: pBCSK-7; site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	65.1%; Score 908; DB 9; Length 963;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches	908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	404 ACAGCGCATGTCTAGTGTCCACTATTCGACCTACCGAACAAGATTAAAAATTAACC	463			
DB	1 ACCAGCGCATGTCTAGTGTCCACTATTCGACCTACCGAACAAGATTAAAAATTAACC	60			


```
QY 464 AAGTAACATATCCAGTGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 523
Db 61 AAGTAACATATCCAGTGAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACC 120
QY 524 AGCATCCTCTTAAACAAAGCATATTTGCAAGAGACAAATTAAGTTACAGTTTACAA 583
Db 121 AGCATCCTCTTAAACAAAGCATATTTGCAAGAGACAAATTAAGTTACAGTTTACAA 180
QY 584 CATCTAAGGCGCAATATATATGAAAGTAGCTATGACGTTGAGATTTTCTTTTC 643
Db 181 CATCTAAGGCGCAATATATATGAAAGTAGCTATGACGTTGAGATTTTCTTTTC 240
QY 644 ATCTGTTATTTTGTATGTTTATATACATTTTCTCTTACAAATAGAGATTT 703
Db 241 ATCTGTTATTTTGTATGTTTATATACATTTTCTCTTACAAATAGAGATTT 300
QY 704 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTCGAG 763
Db 301 TCTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTCGAG 360
QY 764 ATCTGTTCAAAATCTTCTCATTTTATTAAGCTAGTTGGCAACCTGTTCTTT 823
Db 361 ATCTGTTCAAAATCTTCTCATTTTATTAAGCTAGTTGGCAACCTGTTCTTT 420
QY 824 CAAGAATTTTGAATTTTCAAAAAAATTAAGTTATTTTCTCTTATATAAGAAAC 883
Db 421 CAAGAATTTTGAATTTTCAAAAAAATTAAGTTATTTTCTCTTATATAAGAAAC 480
QY 884 ACTTGAATAATGAGTTGCGACAGTACGCTTGAATGTTTTCCATATAATTAACAATCA 943
Db 481 ACTTGAATAATGAGTTGCGACAGTACGCTTGAATGTTTTCCATATAATTAACAATCA 540
QY 944 CTGTGATATATTTTGGCCAGGCCCAATAATTTTAAACGGAACGAACTGAGCGCA 1003
Db 541 CTGTGATATATTTTGGCCAGGCCCAATAATTTTAAACGGAACGAACTGAGCGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAAATCAG 1063
Db 601 AACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGAGAGAGAGAGAGAGAAATCAG 660
QY 1064 TTTTAAGCTATTTCCCTGAGATGTCGGGTTTGGCAACGATGACCAACCGTATCATAGCT 1123
Db 661 TTTTAAGCTATTTCCCTGAGATGTCGGGTTTGGCAACGATGACCAACCGTATCATAGCT 720
QY 1124 CATAGTGCTAGTCAAGGTTGGGCAAGCTCTGTCATCTGCAATGAGGATCTATATG 1183
Db 721 CATAGTGCTAGTCAAGGTTGGGCAAGCTCTGTCATCTGCAATGAGGATCTATATG 780
QY 1184 TTGTTCAACGTTGCTCTGTTTCATGTCGTCAGCAAGCCTTGCTATTTCTGAACCAAGAGAT 1243
Db 781 TTGTTCAACGTTGCTCTGTTTCATGTCGTCAGCAAGCCTTGCTATTTCTGAACCAAGAGAT 840
QY 1244 ACTTACTCCCAAAACATTCATCTTACTCATGCAACTTCGATGCAACGCAATATGTT 1303
Db 841 ACTTACTCCCAAAACATTCATCTTACTCATGCAACTTCGATGCAACGCAATATGTT 900
QY 1304 TCCTGAAC 1311
Db 901 TCCTGAAC 908

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
DEFINITION OGIAG08TV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMMa071B515,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 915)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, O.A., Rohlfing, T.,
Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGIAG08TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Classes: sheared ends.
FEATURES
source
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMMa071B15"
/clone_1ib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 47.8%; Score 667; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 645 TTCTGTTATTTTGTATGTTTATATACATTTTCTCTCTTCAATAGAGATTTT 704
Db 24 TTCTGTTATTTTGTATGTTTATATACATTTTCTCTCTTCAATAGAGATTTT 83
QY 705 CTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTCGAG 764
Db 84 CTTCGATTTTAAATGACTATATAAGTCAATTTTATATAAGACGACGATGTCGAG 143
QY 765 TTCTGTTCAAAATCTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTC 824
Db 144 TTCTGTTCAAAATCTTCTGATTTTAAAGCTAGTTTGGCAACCTGTTCTTTC 203
QY 825 AAAGATTTTGAATTTTCAAAAAAATTAAGTTATTTCTCTTATATAAGAAACA 884
Db 204 AAAGATTTTGAATTTTCAAAAAAATTAAGTTATTTCTCTTATATAAGAAACA 263
QY 885 CTTGAAAAATGAGTTGCCAGACTAGCCCTGAGATGTTTTCCATATAATTAACATC 944
Db 264 CTTGAAAAATGAGTTGCCAGACTAGCCCTGAGATGTTTTCCATATAATTAACATC 323
QY 945 TGTGTATATATTTTGGCCAGGCCCAATAATTTTAAACGGAACGAACTGAGCGGA 1004
Db 324 TGTGTATATATTTTGGCCAGGCCCAATAATTTTAAACGGAACGAACTGAGCGGA 383
QY 1005 ACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGGAGAGAGAGAGAGAAATCACT 1064
Db 384 ACCAATCTGAGCTATTTCTCTAGATTAAGTAAAGGAGAGAGAGAGAGAAATCACT 443
QY 1065 TTTAAGTCAATGCTCCTGAGATGTCGGGTTTGGCAACGATAGCCACCGTATCATAGCT 1124
Db 444 TTTAAGTCAATGCTCCTGAGATGTCGGGTTTGGCAACGATAGCCACCGTATCATAGCT 503
QY 1125 ATAGTGCTTACGTCAGGTTGGGCAAGCTCTGTCATCTCAATGCGATATCATAGCT 1184
Db 504 ATAGTGCTTACGTCAGGTTGGGCAAGCTCTGTCATCTCAATGCGATATCATAGCT 563
QY 1185 TGTTCACCGTTGCTGTTTCCATGTCGCAAGCCTTGCTATTTCTGAACCAAGAGATA 1244
Db 564 TGTTCACCGTTGCTGTTTCCATGTCGCAAGCCTTGCTATTTCTGAACCAAGAGATA 623
```


QY 1245 CCTACTCCCAACAATCATCTTACTGATGCAATCTTCATGCAACGACGACATATGTTT 1304
|
|
|
Db 624 CCTACTCCCAACAATCATCTTACTGATGCAATCTTCATGCAACGACGACATATGTTT 683

QY 1305 CCTGAAC 1311
|
|
|
Db 684 CCTGAAC 690

RESULT 3
CC656939 687 bp DNA linear GSS 19-JUN-2003
LOCUS CG224211/c
DEFINITION OGM20TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBma0554D15,
genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 687)
Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Unpublished (2002)
Consortium for Maize Genomics
Other GSSs: OGM20TV
Contact: Cathy Whitefaw
TIGR

TITLE
JOURNAL
COMMENT
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_id="ZM_0.7-1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 29.3%; Score 409; DB 9; Length 687;
Best Local Similarity 100.0%; Pred.No.1.1e-197;
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 903 CCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCACTGTGTATTAATTTGGC 962
|
|
|
Db 673 CCAGACTAGCCCTAGATGTTTCCCAATTAATTAATCACTGTGTATTAATTTGGC 614

QY 963 CAGCCCCATTAATTTTAAACCGAAATGAAATCGAGCAACCAATCTGAATTTT 1022
|
|
|
Db 613 CAGCCCCATTAATTTTAAACCGAAATGAAATCGAGCAACCAATCTGAATTTT 554

QY 1023 CTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTCCCG 1082
|
|
|
Db 553 CTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTCCCG 494

QY 1083 AGATGTGCGGTTGGCAACGATAGCAGCCGTAATCATAGTCTAGAGTGGCTTACG 1142
|
|
|
Db 493 AGATGTGCGGTTGGCAACGATAGCAGCCGTAATCATAGTCTAGAGTGGCTTACG 434

QY 1143 TTGCGAGCTCTGTGTCTATCTCAATGAGATCTAATGCTTGTCAACCGTTGCTTT 1202
|
|
|
Db 433 TTGCGAGCTCTGTGTCTATCTCAATGAGATCTAATGCTTGTCAACCGTTGCTTT 374

QY 1203 GTTCATGCGCAAGCCCTTCCATTCGATCAACCAAGAGATACCTACTCCCAACAATCC 1262
|
|
|
Db 373 GTTCATGCGCAAGCCCTTCCATTCGATCAACCAAGAGATACCTACTCCCAACAATCC 314

QY 1263 ATCTTACTGCAACTTCATGCAACCAAGAGATATGTTTCTGAAC 1311
|
|
|
Db 313 ATCTTACTGCAACTTCATGCAACCAAGAGATATGTTTCTGAAC 265

RESULT 4
CG224211 878 bp DNA linear GSS 22-AUG-2003
LOCUS CG224211/c
DEFINITION OGIAG08TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBma0716B15,
genomic survey sequence.
ACCESSION CG224211
VERSION CG224211.1 GI:34124099
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 878)
Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Unpublished (2002)
Consortium for Maize Genomics
Other GSSs: OGIAG08TV
Contact: Cathy Whitefaw
TIGR

TITLE
JOURNAL
COMMENT
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..878
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_id="ZMMBma0716B15"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 4.7%; Score 66; DB 9; Length 878;
Best Local Similarity 100.0%; Pred.No.5.5e-22;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1324 GATCAACAAGCTAGGCTTCGCGCTAGCTCCCTGCTCGGCGATCTTTTGGT 1383
|
|
|
Db 878 GATCAACAAGCTAGGCTTCGCGCTAGCTCCCTGCTCGGCGATCTTTTGGT 819

QY 1384 CCACCA 1389
|
|
|
Db 818 CCACCA 813

RESULT 5
BZ828689 539 bp DNA linear GSS 18-MAR-2003
LOCUS BZ828689/c
DEFINITION PUFH74TB ZM 0.6-1.0 KB Zea mays genomic clone ZMMBta320N03,
genomic survey sequence.
ACCESSION BZ828689
VERSION BZ828689.1 GI:29050496
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;


```

REFERENCE
  TITLE      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
  JOURNAL    clade: Panicoideae; Andropogoneae; Zea.
  COMMENT    1 (bases 1 to 539)
              WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
              Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
              Bennetzen,J.
              Maize Genomics Consortium
              Unpublished (2003)
              Other GSSs: PUFHM74TD
              Contact: Cathy WhiteJaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whiteJaw@tigr.org
              Seq primer: TR
              Claes: sheared ends.
FEATURES
  source     Location/Qualifiers
              1..539
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone_lib="ZMBRta320N03"
              /clone_1lb="ZM 0.6 1.0 KB"
              /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"
ORIGIN
  Query Match      1.9%; Score 27; DB 8; Length 539;
  Best Local Similarity 100.0%; Pred. No. 0.052;
  Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      843 CAAAAAAATTAGTTATTTCTCTTT 869
        |||||||
        292 CAAAAAAATTAGTTATTTCTCTTT 266
RESULT 6
LOCUS      BZ828694                611 bp    DNA    linear    GSS 18-MAR-2003
DEFINITION PUFHM74TD_ZM_0.6_1.0_KB Zea mays genomic clone ZMBRta320N03,
ACCESSION  BZ828694
VERSION    BZ828694.1 GI:29050508
KEYWORDS  GSS.
SOURCE     Zea mays
           Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade: Panicoideae; Andropogoneae; Zea.
           1 (bases 1 to 611)
           WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
           Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
           Bennetzen,J.
           Maize Genomics Consortium
           Unpublished (2003)
           Other GSSs: PUFHM74TD
           Contact: Cathy WhiteJaw
           TIGR
           9712 Medical Center Drive, Rockville, MD 20850, USA
           Tel: 301-838-5843
           Fax: 301-838-0208
           Email: whiteJaw@tigr.org
           Seq primer: TR
           Claes: sheared ends.
FEATURES
  source     Location/Qualifiers
              1..611
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone_lib="ZMBRta320N03"
              /clone_1lb="ZM 0.6 1.0 KB"
              /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"
ORIGIN
  Query Match      1.9%; Score 27; DB 8; Length 539;
  Best Local Similarity 100.0%; Pred. No. 0.052;
  Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      843 CAAAAAAATTAGTTATTTCTCTTT 869
        |||||||
        292 CAAAAAAATTAGTTATTTCTCTTT 266

```

```

ORIGIN
  Query Match      1.9%; Score 27; DB 8; Length 611;
  Best Local Similarity 100.0%; Pred. No. 0.052;
  Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      843 CAAAAAAATTAGTTATTTCTCTTT 869
        |||||||
        222 CAAAAAAATTAGTTATTTCTCTTT 248
RESULT 7
LOCUS      CG148339                804 bp    DNA    linear    GSS 21-AUG-2003
DEFINITION PUKB46TB_ZM_0.6_1.0_KB Zea mays genomic clone ZMBRta0780G19,
ACCESSION  CG148339
VERSION    CG148339.1 GI:34039122
KEYWORDS  GSS.
SOURCE     Zea mays
           Zea mays
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade: Panicoideae; Andropogoneae; Zea.
           1 (bases 1 to 804)
           WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
           Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
           Bennetzen,J.
           Maize Genomics Consortium
           Unpublished (2003)
           Other GSSs: PUKB46TD
           Contact: Cathy WhiteJaw
           TIGR
           9712 Medical Center Drive, Rockville, MD 20850, USA
           Tel: 301-838-5843
           Fax: 301-838-0208
           Email: whiteJaw@tigr.org
           Seq primer: TR
           Claes: sheared ends.
FEATURES
  source     Location/Qualifiers
              1..804
              /organism="Zea mays"
              /mol_type="genomic DNA"
              /strain="B73"
              /db_xref="taxon:4577"
              /clone_lib="ZMBRta0780G19"
              /clone_1lb="ZM 0.6 1.0 KB"
              /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
              Cot selected genomic DNA library"
ORIGIN
  Query Match      1.9%; Score 27; DB 9; Length 804;
  Best Local Similarity 100.0%; Pred. No. 0.052;
  Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      837 TTTTTCACAAAAATTAGTTATTTT 863
        |||||||
        156 TTTTTCACAAAAATTAGTTATTTT 182
RESULT 8
LOCUS      CG011564/c                854 bp    DNA    linear    GSS 19-AUG-2003
DEFINITION ZUABT23TV_ZM_3.0_4.0_KB Zea mays genomic clone ZMBRPa0018C22,
ACCESSION  CG011564
VERSION    CG011564.1 GI:33883730
KEYWORDS  GSS.
SOURCE     Zea mays
           Zea mays
           ORGANISM
           Zea mays

```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 854)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Other_GSSs: Z048T23TH
Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@cigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..854

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMPa0018C22"

/note="Vector: pBCK-; Site_1: HincII; 3-4 kb 'unfiltered' genomic DNA library"

ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 854;

Best Local Similarity 100.0%; Pred. No. 0.17; Mismatches 0; Indels 0; Gaps 0;

QY 268 GGGGATATATGTCCTCCACATGTC 293
221 GGGGATATATGTCCTCCACATGTC 196

RESULT 9
BZ371637/c 687 bp DNA linear GSS 26-NOV-2002
LOCUS 1e38n03.5', genomic survey sequence.
DEFINITION
BZ371637
BZ371637
BZ371637.1 GI:25455537

ACCESSION
BZ371637
BZ371637
BZ371637.1 GI:25455537

KEYWORDS
GSS.

SOURCE
Zea mays

ORGANISM
Zea mays

REFERENCE
AUTHORS

Rabinowicz, P.D., O'Shaughnessy, A.L., Ballja, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascimben, L., Zuber, T., McCombie, W.R. and Mariensen, R.A.

Genomic shotgun sequences from Zea mays (methy1-filtered)

Unpublished (2002)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100 Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: 1e38 row: h column: 03

Seq primer: -21M13unlvwd

Class: shotgun

High quality sequence stop: 687.

Location/Qualifiers

1..687

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"
/db_xref="taxon:4577"
/clone="1e38n03"
/lab_host="DH5a"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (x/y reads in M13mp19, b/g reads in pUC19). The same ligation was transformed into DH5a."

ORIGIN

Query Match 1.8%; Score 25; DB 8; Length 687;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAATTAGTTATTTCTCTT 869
363 AAAAAATTAGTTATTTCTCTT 339

RESULT 10
CC810403/c 726 bp DNA linear GSS 16-JUL-2003
LOCUS ZMBC0479C04r ZMBC Zea mays genomic clone ZMBC0479C04 3',
DEFINITION
CC810403
CC810403
CC810403.1 GI:32817028

ACCESSION
CC810403
CC810403
CC810403.1 GI:32817028

KEYWORDS
GSS.

SOURCE
Zea mays

ORGANISM
Zea mays

REFERENCE
AUTHORS

Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C., Rouzard, K., Fuks, G., Yu, Y., Wang, R. and Messing, J.

Sequencing of the maize genome at FGI (2003b)

Unpublished (2003)

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@waksman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 397.

Location/Qualifiers

1..726

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMBC0479C04"

/lab_host="E. coli DH10B"

/clone_lib="ZMBC"

/note="Vector: pTARACT.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 726;

Best Local Similarity 100.0%; Pred. No. 0.55; Mismatches 0; Indels 0; Gaps 0;

QY 265 CCAGGGATATATGTCCTCCACAT 289
364 CCAGGGATATATGTCCTCCACAT 340


```

RESULT 11
CG305519/c      819 bp      DNA      linear      GSS 25-AUG-2003
LOCUS           CG305519
DEFINITION      CG305519 ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0713C15,
                 genomic survey sequence.
ACCESSION       CG305519
VERSION         CG305519.1
KEYWORDS        GI:34219733
SOURCE          Zea mays
ORGANISM        Zea mays
REFERENCE       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                 clade; Panicoidae; Andropogoneae; Zea.
                 1 (bases 1 to 819)
Whitelew,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: CG305519
Contact: Cathy Whitelew
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@tigr.org
Seq primer: TF
Classes: sheared ends.
FEATURES
source
1..819
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 819;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 844 AAAAAAATTAGTTATTTCTCTT 868
Db 142 AAAAAAATTAGTTATTTCTCTT 118

RESULT 12
CG190606/c      825 bp      DNA      linear      GSS 21-AUG-2003
LOCUS           CG190606
DEFINITION      CG190606 ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0585M20,
                 genomic survey sequence.
ACCESSION       CG190606
VERSION         CG190606.1
KEYWORDS        GI:34081670
SOURCE          Zea mays
ORGANISM        Zea mays
REFERENCE       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                 clade; Panicoidae; Andropogoneae; Zea.
                 1 (bases 1 to 825)
Whitelew,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHP82TD
Contact: Cathy Whitelew
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@tigr.org
Seq primer: TF
Classes: sheared ends.
FEATURES
source
1..825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZM_0.6_1.0_KB"
/notes="Vector: pCR4-NOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 845 AAAAAAATTAGTTATTTCTCTT 869
Db 582 AAAAAAATTAGTTATTTCTCTT 606

TITLE
JOURNAL
COMMENT
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

```

```

Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@tigr.org
Seq primer: TF
Classes: sheared ends.
FEATURES
source
1..825
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMMBTA0585M20"
/notes="Vector: pCR4-NOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 825;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 845 AAAAAAATTAGTTATTTCTCTT 869
Db 715 AAAAAAATTAGTTATTTCTCTT 691

RESULT 13
CG190609
LOCUS           CG190609
DEFINITION      CG190609 ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0585M20,
                 genomic survey sequence.
ACCESSION       CG190609
VERSION         CG190609.1
KEYWORDS        GI:34081670
SOURCE          Zea mays
ORGANISM        Zea mays
REFERENCE       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                 clade; Panicoidae; Andropogoneae; Zea.
                 1 (bases 1 to 833)
Whitelew,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Reinick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHP82TD
Contact: Cathy Whitelew
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelew@tigr.org
Seq primer: TF
Classes: sheared ends.
FEATURES
source
1..833
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZM_0.6_1.0_KB"
/notes="Vector: pCR4-NOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 833;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 845 AAAAAAATTAGTTATTTCTCTT 869
Db 582 AAAAAAATTAGTTATTTCTCTT 606

TITLE
JOURNAL
COMMENT
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

```



```

RESULT 14
CC358144/c      857 bp   DNA      linear   GSS 16-MAY-2003
LOCUS           PUHPY12TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBtra507B23,
DEFINITION      genomic survey sequence.
ACCESSION       CC358144
VERSION         CC358144.1   GI:30827544
KEYWORDS        GSS.
SOURCE          Zea mays
ORGANISM        Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
                1 (bases 1 to 857)
REFERENCE       1 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
AUTHORS         Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
                Bennetzen,J.
TITLE           Maize Genomic Consortium
JOURNAL         Unpublished (2003)
COMMENT         Other GSSs: PUHPY12TB
                Contact: Cathy WhiteLaw
                TIGR
                9712 Medical Center Drive, Rockville, MD 20850, USA
                Tel: 301-838-5843
                Fax: 301-838-0208
                Email: whiteLaw@cigr.org
                Seq primer: TP
                Class: sheared ends.
FEATURES
source          Location/Qualifiers
                1..857
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBtra507B23"
                /clone_1lb="ZM_0.6_1.0_KB"
                /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
                Cot selected genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 8; Length 857;
Best Local Similarity 100.0%; Pred.No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 845 AAAAAAATTAGTTATTTCTCTT 869
      |||||||||||||||||||
Db 853 AAAAAAATTAGTTATTTCTCTT 829

RESULT 15
CG307825/c      878 bp   DNA      linear   GSS 26-AUG-2003
LOCUS           OCMBG01TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0527A01,
DEFINITION      genomic survey sequence.
ACCESSION       CG307825
VERSION         CG307825.1   GI:34224985
KEYWORDS        GSS.
SOURCE          Zea mays
ORGANISM        Zea mays
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
                clade; Panicoideae; Andropogoneae; Zea.
                1 (bases 1 to 878)
REFERENCE       1 WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
AUTHORS         Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
                Citek,R.W., Nunez,A., Robbins,D. and Lakey,N.
                Consortium for Maize Genomics
                Unpublished (2002)
                Other GSSs: OCMBG01TV
                Contact: Cathy WhiteLaw
                TIGR

```

```

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@cigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
source          Location/Qualifiers
                1..878
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMBma0527A01"
                /clone_1lb="ZM_0.7_1.5_KB"
                /note="Vector: pBSK+; Site_1: HincII; 0.7-1.5 kb
                methylation filtered genomic DNA library"
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 878;
Best Local Similarity 100.0%; Pred.No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 844 AAAAAAATTAGTTATTTCTCTT 868
      |||||||||||||||||||
Db 167 AAAAAAATTAGTTATTTCTCTT 143

Search completed: September 16, 2005, 17:48:56
Job time : 4293.5 secs

```


This Page Blank (uspto)

QY 361 GGGGCTCAGCATGATTTTGGGCCCCAGAAATGCGAGGACGACGATGCTTAGTG 420
 DB 361 GGGGCTCAGCATGATTTTGGGCCCCAGAAATGCGAGGACGACGATGCTTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAATCAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAATCAATCCACT 480
 QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCTTTAAAAAA 540
 QY 541 CAAGCATATTTCGAAAGAGACAAATTTGTTACAGTTTACAAACATCTTAAGACGACAA 600
 DB 541 CAAGCATATTTCGAAAGAGACAAATTTGTTACAGTTTACAAACATCTTAAGACGACAA 600
 QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTAATTTGTT 660
 DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTAATTTGTT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGAGTATTTTCCGATTTTATATA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGAGTATTTTCCGATTTTATATA 720
 QY 721 ATGACATATAAGTCATTTTATATATAGACGACGATGCTAGATTTCTGTTCAAAAATC 780
 DB 721 ATGACATATAAGTCATTTTATATATAGACGACGATGCTAGATTTCTGTTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTTGAATTT 840
 DB 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTTGAATTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTTCTTTTATATAATAGAAACATTTAGAAAAATAGAGT 900
 DB 841 TTCAAAAAAATTAAGTTATTTTCTTTTATATAATAGAAACATTTAGAAAAATAGAGT 900
 QY 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGCTGTAATTAATTTG 960
 DB 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGCTGTAATTAATTTG 960
 QY 961 GCCAGCCCCATTAATTTTAAACCGAACTGAAGTGAAGGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCCATTAATTTTAAACCGAACTGAAGTGAAGGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTTAGATTAAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
 DB 1021 TTCTTAGATTAAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
 QY 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGCTATATAGCTCATAGTGCTTACGTCA 1140
 DB 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGCTATATAGCTCATAGTGCTTACGTCA 1140
 QY 1141 GGTTCGGCAGCTCTCGTGTCATCTCAATGAGCACTACTGCTTTCAACCGTTCGTC 1200
 DB 1141 GGTTCGGCAGCTCTCGTGTCATCTCAATGAGCACTACTGCTTTCAACCGTTCGTC 1200
 QY 1201 TTGTTCCATGTCGCAAGCCTTGCTTATCTGAACCAAGAGATACCTATCCCAAAAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCTTGCTTATCTGAACCAAGAGATACCTATCCCAAAAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCACTATGTTTCTGAAACAGATCTATT 1320
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCACTATGTTTCTGAAACAGATCTATT 1320
 QY 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 DB 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 QY 1381 CGTCCACACCATG 1394
 DB 1381 CGTCCACACCATG 1394

RESULT 3
 AX224394 1394 bp DNA linear PAT 10-SEP-2001
 LOCUS AX224394
 DEFINITION Sequence 1 from Patent WO0160997.
 ACCESSION AX224394
 VERSION AX224394.1 GI:15554636
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays
 Zea mays
 Bukariyoca; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 REFERENCE
 1 Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
 Male tissue-preferred regulatory region and method of using same
 Patent: WO 0160997-A 1 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 Location/Qualifiers
 source
 1..1394
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"
 ORIGIN
 Query Match 99.7%; Score 1389.2; DB 6; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 3,8e-268;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CCATGGTGCTCTATGAAAAAGATGATCAATGTGCTATATCGTTTCTTAGGGTCC 60
 DB 1 CCATGGTGCTCTATGAAAAAGATGATGATCAATGTGCTATATCGTTTCTTAGGGTCC 60
 QY 61 CTTCCTCTGCTTATTAAGTACTGAATCGGGTTTCAAAAACTTCCAGGGTGCATGAT 120
 DB 61 CTTCCTCTGCTTATTAAGTACTGAATCGGGTTTCAAAAACTTCCAGGGTGCATGAT 120
 QY 121 CTCGATGTTCCACTTTCACACCTGCGGTTGCACATTTCTTGATGTCGGTGTCCAT 180
 DB 121 CTCGATGTTCCACTTTCACACCTGCGGTTGCACATTTCTTGATGTCGGTGTCCAT 180
 QY 181 CTGACCGAGGCCCATGACGACCTTTCGAGACCCCATCAAGGCTTTGCGATGCGCCA 240
 DB 181 CTGACCGAGGCCCATGACGACCTTTCGAGACCCCATCAAGGCTTTGCGATGCGCCA 240
 QY 241 CGAAGCATTCGGGTGCTGATGATCCAGGGATATATGTCGCCCAATGCTCACTATA 300
 DB 241 CGAAGCATTCGGGTGCTGATGATCCAGGGATATATGTCGCCCAATGCTCACTATA 300
 QY 301 TTATTAATCTTATGATTAATTTTATTTTGGAAAAATAACAATTAATTAATCTTTGTTA 360
 DB 301 TTATTAATCTTATGATTAATTTTATTTTGGAAAAATAACAATTAATTAATCTTTGTTA 360
 QY 361 GGGGCTCAGCATGATTTTGGGCCCCAGAAATGCGAGGACGACGATGCTTAGTG 420
 DB 361 GGGGCTCAGCATGATTTTGGGCCCCAGAAATGCGAGGACGACGATGCTTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAATCAATCCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATAACCAAGTAATCAATCCACT 480
 QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCTTTAAAAAA 540
 DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGATCTTTAAAAAA 540
 QY 541 CAAGCATATTTCGAAAGAGACAAATTTGTTACAGTTTACAAACATCTTAAGACGACAA 600
 DB 541 CAAGCATATTTCGAAAGAGACAAATTTGTTACAGTTTACAAACATCTTAAGACGACAA 600
 QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTAATTTGTT 660
 DB 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCATCTTGTAATTTGTT 660


```

Qy 661 ATTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
Db 661 ATTGTTTTATATACATTTTCTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
Qy 721 ATGACTATAAAGTCATTTTATATATAGAGACCGATGCTGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATATAGAGACCGATGCTGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTATAGAGCTAGTTTGGCAACCGTGTCTTCAAAAATTTGATTTT 840
Db 781 TTTCTGATTTTATAGAGCTAGTTTGGCAACCGTGTCTTCAAAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTATTTTCTTTATATAATAGAAAACATTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATAGTTATTTTCTTTATATAATAGAAAACATTTAGAAAAATAGAGT 900
Qy 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTCATCTGTTGTAATTAATTTTGG 960
Db 901 TGGCAGACTAGCCCTAGATGTTTCCCAATTAATTCATCTGTTGTAATTAATTTTGG 960
Qy 961 GCCAGCCCATTAATTTATTAACCGAAATCGAAATCGAGGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCATTAATTTATTAACCGAAATCGAAATCGAGGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAACTCATTTGCC 1080
Qy 1081 TGAAGATGTCGGTTTGGCAACGATAGCCATGATGCTCATAGTGGCTTACTGCTA 1140
Db 1081 TGAAGATGTCGGTTTGGCAACGATAGCCATGATGCTCATAGTGGCTTACTGCTA 1140
Qy 1141 GGTTCGGAGCTCTCGTGTATCTACATGAGCAATGCTTGTTCACACCGTTCGTC 1200
Db 1141 GGTTCGGAGCTCTCGTGTATCTACATGAGCAATGCTTGTTCACACCGTTCGTC 1200
Qy 1201 TTGTTCCATGCTCAAGCCCTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Db 1201 TTGTTCCATGCTCAAGCCCTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Qy 1261 CCATCTTATCTAGCAACTTCCATGCAAAACGACATATGTTTCTGGAACAGATCTATT 1320
Db 1261 CCATCTTATCTAGCAACTTCCATGCAAAACGACATATGTTTCTGGAACAGATCTATT 1320
Qy 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGGCTTCCCTCTCTGCTGCGATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGGCTTCCCTCTCTGCTGCGATCTTTT 1380
Qy 1381 CGTCCACCAACCATG 1394
Db 1381 CGTCCACCAACCATG 1394

```

```

RESULT 4          1394 bp      DNA      linear      PAT 27-AUG-2002
BD062176          Male tissue-preferred regulatory region and method of using same.
DEFINITION       BD062176
ACCESSION        BD062176
VERSION          BD062176.1 GI:22607781
KEYWORDS         JP 2001520523-A/1.
SOURCE           Homo sapiens (human)
ORGANISM         Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE        1 (bases 1 to 1394)
AUTHORS          Albertsen,M.C., Fox,T.W., Garnaat,C.W., Hufman,G.A. and
                  Kendell,T.L.
TITLE            Male tissue-preferred regulatory region and method of using same
JOURNAL          PIONEER HI BRED INTERNATIONAL INC
COMMENT          PN JP 2001520523-A/1
                  PD 30-OCT-2001
                  PF 19-JUN-1998 JP 199504910

```

```

PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAT, GARY A HUFMAN,
PI TERRY L KENDALL.
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
, C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 1389.2; DB 6; Length 1394;
Best Local Similarity 99.8%; Pred. No. 3.8e-268;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCATGTCCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGGGTCC 60
Db 1 CCATGTCCTCTATGAAAAAGATGATGATGCTATATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTTCTGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Db 61 CTTCTTCTGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTGCGTTGCACTTTCTTGGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTGCGTTGCACTTTCTTGGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACTTTCCGGAACCCATCAAGGCGCTTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACTTTCCGGAACCCATCAAGGCGCTTTGGATGGCCCA 240
Qy 241 CGAAGCATATCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Db 241 CGAAGCATATCGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
Qy 301 TTATTAATCTTATGATATTTATTTTGGAAAAATTAACAACCTTATCTTTTGTGTA 360
Db 301 TTATTAATCTTATGATATTTATTTTGGAAAAATTAACAACCTTATCTTTTGTGTA 360
Qy 361 GGGGCTCAGCATATGTTTTCGTTAGGGCCGAAATGCGAGACCAAGCATGCTAGTG 420
Db 361 GGGGCTCAGCATATGTTTTCGTTAGGGCCGAAATGCGAGACCAAGCATGCTAGTG 420
Qy 421 TCCACTATTTGGCATACCCAGAACAGATTTAAAAAATTAACCAAGTAACTTAATCACT 480
Db 421 TCCACTATTTGGCATACCCAGAACAGATTTAAAAAATTAACCAAGTAACTTAATCACT 480
Qy 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTTTAAACCAAGATCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTTTAAACCAAGATCTCTTAAAAAA 540
Qy 541 CAAGATATTTGAAAGAGACAAATTTATGTTTACATTTTAAACCAAGTAAAGCGACAAA 600
Db 541 CAAGATATTTGAAAGAGACAAATTTATGTTTACATTTTAAACCAAGTAAAGCGACAAA 600
Qy 601 TTATATCGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 660
Db 601 TTATATCGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 660
Qy 661 ATTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
Db 661 ATTGTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
Qy 721 ATGACTATAAAGTCATTTTATATATAGAGACCGATGCTGATTTCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATATAGAGACCGATGCTGATTTCTGTTCAAAAATC 780
Qy 781 TTTCTGATTTTATAGAGCTAGTTTGGCAACCGTGTCTTCAAAAATTTGATTTT 840

```


Db	781	TTTTTGATTTTTTTAAAGACTAGTTGGCAACCCTGTTCTTTCAAAGAAATTTGATTTT	840
Oy	841	TTCAAAAAAATTAGTTTATTTTCTCTTTATTAATAATAGAAAAACATTAGAAAAATAGAGT	900
Db	841	TTCAAAAAAATTAGTTTATTTTCTCTTTATTAATAATAGAAAAACATTAGAAAAATAGAGT	900
Oy	901	TGCCAGACTAGCCCTAGAAATGTTTTCCCATTAATAATACATCATCTGTATTAATATTG	960
Db	901	TGCCAGACTAGCCCTAGAAATGTTTTCCCATTAATAATACATCATCTGTATTAATATTG	960
Oy	961	GCCAGCCCATTAATTTATTTAAACCGAAATCTGAATTCGAGGAAACCAATCTGAGCTAT	1020
Db	961	GCCAGCCCATTAATTTATTTAAACCGAAATCTGAATTCGAGGAAACCAATCTGAGCTAT	1020
Oy	1021	TTCTCTAGATTAGTAATAAAGGAGAGAGAGAAATCATGTTTAAATCATTTGTC	1080
Db	1021	TTCTCTAGATTAGTAATAAAGGAGAGAGAGAAATCATGTTTAAATCATTTGTC	1080
Oy	1081	TGAGATGTCGGTGTGGCAAGATAGCCACCTATTCATAGCTCATAGCTGCTACGTCA	1140
Db	1081	TGAGATGTCGGTGTGGCAAGATAGCCACCTATTCATAGCTCATAGCTGCTACGTCA	1140
Oy	1141	GATTGGAGAGCTCTGTGTCATCTCACATGGCATCTCATGCTTGTTCACACCTGTC	1200
Db	1141	GATTGGAGAGCTCTGTGTCATCTCACATGGCATCTCATGCTTGTTCACACCTGTC	1200
Oy	1201	TTGTTTCATCGTCCAAAGCTTGGCTCTATTTCTGAACCAAGAGATACCTACTCCAAACAT	1260
Db	1201	TTGTTTCATCGTCCAAAGCTTGGCTCTATTTCTGAACCAAGAGATACCTACTCCAAACAT	1260
Oy	1261	CCATTTACTCATGCACTTCCTCATCAACACGACATATGTTTCTGAACCAATCCAT	1320
Db	1261	CCATTTACTCATGCACTTCCTCATCAACACGACATATGTTTCTGAACCAATCCAT	1320
Oy	1321	AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCGCCGATCTTTT	1380
Db	1321	AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCGCCGATCTTTT	1380
Oy	1381	CGTCACCAACCATG 1394	
Db	1381	CGTCACCAACCATG 1394	
RESULT 5			
AF360356			
LOCUS	AF360356	3343 bp	DNA linear
DEFINITION	Zea mays male fertility protein (Ms45) gene, complete cds.		
ACCESSION	AF360356		
VERSION	AF360356.1	GI:14028756	
KEYWORDS			
SOURCE			
ORGANISM			
	Zea mays		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
	1 (bases 1 to 3343)		
	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
	Cloning of Ms45, a gene required for male fertility from Zea mays		
	Unpublished		
	2 (bases 1 to 3343)		
	Fox,T.W., Trimnell,M.R. and Albertsen,M.C.		
	Direct Submission		
	Submitted (13-MAR-2001) Trait and Technology Development, Pioneer		
	Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,		
	IA 50131-1004, USA		
FEATURES			
source			
	1..3343		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/cultivar="B73"		
	/db_xref="taxon:4577"		
	/chromosome="9L"		

Query Match	Best Local Similarity	99.5%	Score	1387.6	DB	8	Length	3343
Matches	1390	Conservative	0	Mismatches	4	Indels	0	Gaps
QY	1	CCATGCTCTCTATGAAAAAGATGATACAAATGTCCTATATCCGTTTCTTAAGGCTC	60					
DB	1	CCATGCTCTCTATGAAAAAGATGATACAAATGTCCTATATCCGTTTCTTAAGGCTC	60					
QY	61	CTTCTTCGCTTTTACTGACCTAACTGGGGTTACAAAACCTTCCACGGGTGCATAT	120					
DB	61	CTTCTTCGCTTTTACTGACCTAACTGGGGTTACAAAACCTTCCACGGGTGCATAT	120					
QY	121	CTCCATGTTCCACTTCTCCCACTCGCGTTGCACATTTCTTGATGTCGGTGTTCCCAT	180					
DB	121	CTCCATGTTCCACTTCTCCCACTCGCGTTGCACATTTCTTGATGTCGGTGTTCCCAT	180					
QY	181	CTGACCGAGGCCCATCAGACACCTTTGGGACACCCCATCAAGGGCTTTGCGATGGCCCA	240					
DB	181	CTGACCGAGGCCCATCAGACACCTTTGGGACACCCCATCAAGGGCTTTGCGATGGCCCA	240					
QY	241	CGAAGCGATGGGGCGTGGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATTA	300					
DB	241	CGAAGCGATGGGGCGTGGTGATCCAGGGGATATATGTCGCCCAATCGTCACTATTA	300					
QY	301	TTATTTATCTTTAATATATTTATTTATTTTGGAAAAATTAACAATCTTATCTTTTGTGTA	360					
DB	301	TTATTTATCTTTAATATATTTATTTATTTTGGAAAAATTAACAATCTTATCTTTTGTGTA	360					
QY	361	GGGCTCGACATAGATTTTGGCTTAAGGGCCCGAAGATGCGAGACCAAGCATGTCATAGT	420					
DB	361	GGGCTCGACATAGATTTTGGCTTAAGGGCCCGAAGATGCGAGACCAAGCATGTCATAGT	420					
QY	421	TCCACTATGGCACTACCCAGAACAAATTTAAAAAATAACAAAGTAACTAATCCACT	480					
DB	421	TCCACTATGGCACTACCCAGAACAAATTTAAAAAATAACAAAGTAACTAATCCACT	480					
QY	481	CGAAGCATATGATGTAATGTTTAAAGAAACATCTATTAAAAACAAGATCCCTTAAAAAAA	540					
DB	481	CGAAGCATATGATGTAATGTTTAAAGAAACATCTATTAAAAACAAGATCCCTTAAAAAAA	540					
QY	541	CAAGCATATTTGAAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA	600					
DB	541	CAAGCATATTTGAAAAGAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA	600					
QY	601	TTATATGAAAGGTATGATGATGACGTTTCTTTTCAATCTTGTTATTTTGT	660					
DB	601	TTATATGAAAGGTATGATGATGACGTTTCTTTTCAATCTTGTTATTTTGT	660					
QY	661	ATTGTTTTTAAATACATTTTCTCTCTTAACTAGAGGATTTTCTTCCGATTTTATATAA	720					
DB	661	ATTGTTTTTAAATACATTTTCTCTCTTAACTAGAGGATTTTCTTCCGATTTTATATAA	720					

QY 721 ATGACTATAAAGTCATTTTATATTAAGACGACGATGCTAGATTCTGCTCAAAAATC 780
 |||||
 DB 721 ATGACTATAAAGTCATTTTATATTAAGACGACGATGCTAGATTCTGCTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAAAGAAATTTGATTTT 840
 |||||
 DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTTCTCTTTTAAATAAGAAAACATTAAGAAAATAGACT 900
 |||||
 DB 841 TTCAAAAAAATAGTTTATTTTCTCTTTTAAATAAGAAAACATTAAGAAAATAGACT 900
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTATTAATTAATG 960
 |||||
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTATTAATTAATG 960
 QY 961 GCCAGCCCCATAAATTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 |||||
 DB 961 GCCAGCCCCATAAATTATTTAAACGAAACTGAATCGAGCAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATCACTTTTAAGTCAATGCTCC 1080
 |||||
 DB 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAAATCACTTTTAAGTCAATGCTCC 1080
 QY 1081 TGAGATGTGCGGTTTGGCAACGATGCAACCGTAATCATAGCTCATAGGTGCTAGCTCA 1140
 |||||
 DB 1081 TGAGATGTGCGGTTTGGCAACGATGCAACCGTAATCATAGCTCATAGGTGCTAGCTCA 1140
 QY 1141 GGTTCGGCAGCTCTCGTGTATCTCACATGCACTACTAGCTGTTGTTCAACCGTTCGTC 1200
 |||||
 DB 1141 GGTTCGGCAGCTCTCGTGTATCTCACATGCACTACTAGCTGTTGTTCAACCGTTCGTC 1200
 QY 1201 TTGTTTCATGTCGCAACGCTTGCTCTTGAACCAAGAGATTAAGTCTCCCAACAT 1260
 |||||
 DB 1201 TTGTTTCATGTCGCAACGCTTGCTCTTGAACCAAGAGATTAAGTCTCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGGAACAGATCTAT 1320
 |||||
 DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGGAACAGATCTAT 1320
 QY 1321 AAGATTCACACAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 |||||
 DB 1321 AAGATTCACACAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 QY 1381 CGTCCACCAACGATG 1394
 |||||
 DB 1381 CGTCCACCAACGATG 1394

RESULT 6
 AX224402 255 bp DNA linear PART 10-SEP-2001
 LOCUS AX224402
 DEFINITION Sequence 9 from Patent WO0160997.
 ACCESSION AX224402
 VERSION AX224402.1 GI:15554644
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1
 AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL PATENT: WO 0160997-A 9 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1..255
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

ORIGIN

Query Match 17.5%; Score 243.6; DB 6; Length 255;
 Best Local Similarity 98.4%; Pred. No. 1.3e-38;
 Matches 246; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1145 CGGAGCTCTCGTGTATCTCAATGCAATGATGCTTGTCAACCGTTCGTTGT 1204
 |||||
 DB 5 CGGAGATCCGCTGTCATCTCAATGCAATGATGCTTGTCAACCGTTCGTTGT 64
 QY 1205 TCCATCTCCAGCCCTTGCTTCTTGAACCAAGAGATACCTACTTCCCAACATCCAT 1264
 |||||
 DB 65 TCCATCTCCAGCCCTTGCTTCTTGAACCAAGAGATACCTACTTCCCAACATCCAT 124
 QY 1265 CTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGGAACAGATCTATTAAG 1324
 |||||
 DB 125 CTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGGAACAGATCTATTAAG 184
 QY 1325 ATCACAACAGTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTTCGTC 1384
 |||||
 DB 185 ATCACAACAGTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTTTCGTC 244
 QY 1385 CACCAACCATG 1394
 |||||
 DB 245 CACCAACCATG 254
 RESULT 7
 AX224396 158 bp DNA linear PART 10-SEP-2001
 LOCUS AX224396
 DEFINITION Sequence 3 from Patent WO0160997.
 ACCESSION AX224396
 VERSION AX224396.1 GI:1554638
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1
 AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL PATENT: WO 0160997-A 3 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1..158
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

ORIGIN
 Query Match 10.5%; Score 146; DB 6; Length 158;
 Best Local Similarity 99.4%; Pred. No. 4.9e-19;
 Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1155 CGTGTATCTCATGATGCAATGATGCTTGTCAACCGTTCGTC-TTGTTCATGCTC 1213
 |||||
 DB 1 CGTGTATCTCATGATGCAATGATGCTTGTTCACCGTTCGTCCTTGTTCATGCTC 60
 QY 1214 CAAAGCTTGCCATTTCTGAACCAAGAGATACCTACTTCCCAACATCATCTTACTCT 1273
 |||||
 DB 61 CAAAGCTTGCCATTTCTGAACCAAGAGATACCTACTTCCCAACATCATCTTACTCT 120
 QY 1274 GCACTTCATGCAACACGACATATGTTTCTGTAAC 1311
 |||||
 DB 121 GCACTTCATGCAACACGACATATGTTTCTGTAAC 158
 RESULT 8
 AC147602 186199 bp DNA linear HTG 17-AUG-2004
 LOCUS AC147602
 DEFINITION Zea mays clone ZMBRC0334A01, *** SEQUENCING IN PROGRESS ***
 ACCESSION AC147602
 VERSION AC147602.5 GI:51315585

KEYWORDS
SOURCE
ORGANISM

HTG; HTGS_PHASE2; HTGS_PULLTOP; HTGS_ACTIVEFIN

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (baves 1 to 186199)

REFERENCE AUTHORS

Blirren, B., Nuebaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K. and Messing, J.

Zea mays, clone ZM8B8C0334A01
Unpublished

TITLE
JOURNAL
REFERENCE

2 (bases 1 to 186199)

Barron, B., Huebner, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Bama, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukrgalter, B., Camarata, J., Chang, Y., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Detrelano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gargya, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Macdonald, T. R., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mieng, V., Murphy, T., Nguyen, C., Nelson, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkhant, P., Pierre, N., Rachukpa, A., Ramesam, U., Raymond, C., Retts, R., Rise, C., Roman, J., Schauer, S., Schnuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodose, J., Topham, K., Travers, M., Vassiliev, L. H., Venkatarman, V. S., Viet, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zairoun, J., Zemek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE

Direct Submission
Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186199)

AUTHORS Birren, B.; Nusbaum, C.

Meising, J. M., Bouhassira, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. B., Barna, N., Bastien, V., Bloom, T., Boguslavsky, A., Bouhassira, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeAvellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorrie, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyn, S., Graham, J., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karakas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Milenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Nottbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkharn, P., Pierre, N., Rachuganga, A., Ramsammy, U., Raymond, C., Rette, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnuppback, R., Seaman, S., Seaver, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, C., Strubbe, M., Talmes, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataratnam, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M. Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA. On Aug 17, 2004 this sequence version replaced gs14958659. All repeats were identified using RepeatMasker: <http://ftp.genome.washington.edu/RM/RepeatMasker.html> Smit, A.F.A. & Green, P. (1996-1997)

----- Genome Center -----

Center: Whitehead Institute/MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

contact: sequence_submission@broad.mit.edu

.....

Bhargava, AK and Messing, J: The Plant Genome

Rutgers, Waksman Institute, Rutgers, The State University of New

(<http://pgr.rutgers.edu>)

Butler, E and Wing, R: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (<http://www.genome.arizona.edu>)

----- Project Information -----

center project name: L30000

Center clone name: 334_A_1

----- consensus information -----

This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ns are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository (<http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?>). An exact list of reads used in this assembly are available at <http://www.broad.mit.edu/annotation/plants/maize/randomclones.html>

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

*	100617:	config of 100617 bp in length
*	100618:	gap of unknown length
*	100717:	gap of unknown length
*	104730:	config of 4013 bp in length
*	104731:	104730: gap of unknown length
*	104831:	115104: config of 10274 bp in length
*	104831:	115104: config of 10274 bp in length
*	115105:	115204: gap of unknown length
*	115205:	115204: gap of unknown length
*	156386:	config of 41192 bp in length
*	156486:	gap of unknown length
*	179936:	config of 23440 bp in length
*	179937:	gap of unknown length
*	186199:	config of 6163 bp in length.
*	180037:	

FEATURES
source

```

/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBc0334A01"
/clone_lib="CHORI-201 Maize B73 BAC Mbol library"

```

ORIGIN

Query Match	5.5%	Score 77	DB 2	length 186199
Best Local Similarity	71.4%	Pred. No. 2.5e-05		
Matches 115	Conservative 0	Mismatches 45	Indels 1	Gaps 1

Qy	Db
780	162058
TTTTTCGATTTTTTAAGAGCTAFTTGGCAACCGTGTCTTTCAAGAATTTTGATT	CTTTCTTTACTTTTAGGGCTAFTTGGGAACCTATT-TTCTAAGAAATTTTCAATT
839	162116

ACCESSION AC117267 AC115597
VERSION AC117267.2 GI:42733680
KEYWORDS HTG
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
AUTHORS 1 (bases 1 to 25769)
Gloeckner,G., Eichinger,L., Szafrański,K., Pachbat,J., Dear,P.,
Lehman,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
Tungel,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
Noegel,A.A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
REFERENCE 2 (bases 1 to 25769)
AUTHORS Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 25769)
REFERENCE Baumgart,C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from GeneID may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/) and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
Location/Qualifiers
1..25769
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
complement(join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="Geneid exon scores (in order of location ranges) :
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38625.1"
/translation="MRLQIILFLLSTILIFOLKSTPSSNDISSTLISIPNPKVDNNO
ILSKSEMFRLSESSSKKVIFFVDFGFNHYEBNDRIIKERENKVGVDSPFCL
NSKPNYLEVIKKFSTISRNKKVYPENIWMISTHCNNSNYLYNNKNNNSQYSLV
FKQVAKSFLQNGYNIGIKSSKNTFKIFENKDKLFSLPFYLNKIAENGFEDYY
SSPLSGMTDFIKQFKKNDIETILSDVEVNLCSNNNNNNNNVDDFTLLGSGSG
SSSGSGSGTSGAOLSGTGVYINHO"
join(2957..3173,3272..5583)
/note="Geneid exon scores (in order of location ranges) :
24.88, 254.96 - GSCJ_ID dd_00753"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38626.1"
/db_xref="GI:42733682"
/translation="MEFPKYIYITNSLDKPKEMWTFQSMSTLTITTDITQKSPVNSG
SSSSSKGFEAVLDLDTAGQEEYSAMDQVKTGDCPMIVFSIDSRSPFEVSQLK
CHIERKQDDVPIIIVGNKVDLESRRQVRIEADQLASLRVPIETSAKTRSIER
AFETLRHTPRNTVVYKVVNGGQVGSALIIIOFONHVEVEYPTIEDSYKQVVIS
GPIPIGSLNKKSSSSSSSSSGKTGLFNKIPSGKQOPQQAASSTIDRIGOI
STNRLEANTLSYSMSLSKEVPLITGDCYCGCGAYILSRPSNLKATGDSPTMKCEP
CKYSNSNIIIEGELPNKDSVEYVLSPESTSTDSKREBSIIICYCIDVSGSWKITT
EVPSLOSERNNAKGVKASGSPYSIRLECVQSSIPMTIDRLSIOYPMKRVLVTFS
DEVMITYDSNSVDGPVIVAGDKLEDFDQLEIGRSWYDKLPASGSSDPFLAKIKLSL

EPVOSTALGPALLVSAIAOSKMLSEVICTDGVNVLGAIEDLPLGAPQFEYEKVT
KLAONKTTTINIGSGSHIDGVIGXVSEQNNNTIITHLELRERKLTQNPMTA
TDVENSICLHPTLEINKDSDKQSGRVYKQFPNNVSLTDITLISRRNPTFVQIY
FOIQTIKLDGVRCLRVASQLOATQTPNFTSIVANISILMAKTAQOODEY
MESRHLKAAATLIRSLCNTDEQSEIYNFVQIOSSSLSSTSCSTYQSCPIDEHIOE
DEIRVFMKQVHKKFVVEGRKQDISRRKGEAEINKQYNIKFT"
complement(5711..6796)
/note="Geneid exon scores (in order of location ranges) :
68.61 - GSCJ_ID dd_00752"
/codon_start=1
/product="similar to Plaeomodium falciparum (isolate 3D7).
Hypothetical protein"
/protein_id="AAS38627.1"
/db_xref="GI:42733683"
/translation="MSRRVGLIRRYNGINLTIEKSLSPRLINNTNNNNNN
GTITTTTGGANNNSSTPSTPVSPFDVYDHSNNQIOYVDSVYVNNNNNN
NN
KFNNNNNYFPSTNEMITTLNLSCTLEDIGTNDVSTNTPTSPSSSTSLN
KSNHMMPSLKINOIINDNNNISINFEVQIOSSSLSSTSCSTYQSCPIDEHIOE
OLRIQEKQKOIDDODDDDDDDDDSSPPPPPTTTTTPCLSPRPIKKEENID
QNSMPPPKLDRRSPRLIPTTANT"
join(8924..9076,9236..10573)
/note="Geneid exon scores (in order of location ranges) :
10.37, 128.60 - GSCJ_ID dd_00729"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38628.1"
/db_xref="GI:42733684"
/translation="SSSTITTPATTTTNNNSNGIHNSKTNLTFLKNIKNIPNS
PNNAPIKKNVGIIGNGNVNNNNNDKRLNSDNGVFNPKNLTLPSPKSNMSTKHS
SSSSLLAKNEYDIESIKELKSMDSITGGLSPSSPRESPISTLDESLATTK
LTDVAASSSSSVKTYIDVLSFKSANTRPDIIVENHOKIIVPLSLPTPN
NNNNNNNNNSOHTANTAGOKPIEPSQIOPKPIAYQASSTKTNNNNNNTPT
TKKANSITTPQSNRNHNHNSNTKSKQNPPIPSLNTATPNOKNTTPSKSLA
TTPSKNKINDVNAAPAAVANSSTTTEQVVAIIDATPSTNATPKTPLKEPDRLA
ARORIDIGPKTAGYKNYIALVPKSRKPTPKPNKQVCSKRWQOIKKMRQLH
ENYDTGIPNEKEVAESNNESTESTLEKENKLEISEENHNH"
complement(10972..11979)
/note="Geneid exon scores (in order of location ranges) :
23.29 - GSCJ_ID dd_00728"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38629.1"
/db_xref="GI:42733685"
/translation="MEKIKYPRRIIDRONKPKIOEFTSKLIGKESPTTKPIY
EKLAMTOSIESILKNTGTLNKSIFBETSYCGTGFLEBYRPIVYLKTTTBBII
DDYLFETKNSLNMKNEKEKILNSILTYNNNENKIGEFWEIINEFONFTKESFOR
IINNLSWKISIDSRNINSNPINSRCSPFYEFKRSQDAGDFILITSMIGIMONY
ENSILSEKQTIYVHAQSFLLINDLSFNRBIENDLLVYIKLAIOLNSIOLSID
KTYELIDHYLKLFSITLILKQVNDSTYQLKQVRQNSKILSGIYPAHKSXRY
N"
join(12884..13020,13172..15320)
/note="Geneid exon scores (in order of location ranges) :
2.42, 136.56 - GSCJ_ID dd_00727"
/codon_start=1
/product="similar to Dictyostelium discoideum (s)lime
mold). Hypothetical protein"
/protein_id="AAS38630.1"
/db_xref="GI:42733686"
/translation="MDNLIYDFLIRKEYNQOSFLDCILILSINGLYISYDGLDYCN
PRLIENKTVKREALSYGLANTQYQLDEFRLLKLKKRTDCKRIINTHSLAKCC
GNKEDVSIILIIENNNNIINKDOFNNNNNNNNNNNNNNNNNNNNNNNNNNN
TLKNERJINGDGEITTTTTTTTTSRSSTTTTNGVNNNNVLAERQVNRVIR
TOILFHLALNIAHANKKVFPLPIQLVDYKFGYQASVLANYHDNDADSDYSOI
DDDSYSDGNCSSSSSSSDIGSSSNSINNGISSSSSLNSLSLEISMTIRNSL
PEGKSLBEFEKEYVITQDRLPLPSISIRFPGNORLAKVISOVVISITFGSPQ
SLDGWLPKQLOKHIOFGHKFOQTIKQGLPELSLSILDPSYKGVIGSIPDSVT
LDVKNSSQNOESISPNIPKSIPLRVDSFNONIKANDISSNVTLSIHGEHNS
DIGIKLPSNIRLIRKGRAPDRDILCGSSTSIDFGKRNRPVLSMTQTLTSDRGS
KPNQIIPGCIITHKLKSLANGYHPNQIITPADTLPTLESJLNGSTYREIYKNDIED
CYGSLNKGQFSNNSNFCVGTGNNGLREMLKNTSLTTLTLANFNRIEYGDLPNSI
ESLNGYHFNQPIGNVNLPKLLKLFLINSEFNONISADGCI PFGLOTIYIRNSMNF
INSINPLITKYINIIDLSH."


```

CDS
join(16184..16199,17272..17613,17707..19331)
/note="Geneid exon scores (in order of location ranges) :
0.08_32.32_227.42 - GSCU_ID dd_00726"
/codon_start=1
/product="similar to Dictyostelium discoideum (Slime
mold). Spore coat protein Sp87 precursor (Pl3 protein)"
/protein_id="AAS38631.1"
/db_xref="GI:42733687"
/translation="MFIGEDCNKITTNEEBCKKSECIYVNIPECCGHOAKSGKTFD
TCTIENSGYRNSSNNQVVENSNKCFNLDGFIKTIITPEYSGCDMKIKELALGSCSF
QKSSGNSPTSCSPGESICEGLSSGSSGSSGSGSGTSGSSGSSGSGTSGSSGSS
SGSSGSGVSSGCTTTCPEGYHGSMDVAVCLASTTGTTGTLPTSSSTAGVSLCTTT
LCPIGHICVEDNSGNSVCVPCNGGSGTSSGSGTSGGHPDPRDVCDDGPHCEKDDK
TAKVPSPTTSSSSSTSGHDPDPCDVTCTPDGFHCECKDGTAKCVSPPTTGSSTSGG
NTNCSNVNCPDGFYCECKDGTAKCVPSGTOPKCPVSLRCPNHECRFNDQHQ
CCVAVHNDRCRLRPHGHECVNDHNGKCCVRSRHPPEVSLRCPKHECKRDBDGHQ
KKCCVHIHDEVDCLDGRGFECKRIHDSGCVSRHPHPHEKCNKCPGHECK
VDHGEKCCVVAHRPPKCSLRCPHRCRNNHGECCVAVHDKSLRCPHECK
VDHGEKCCVVAHRPPKCSLRCPHRCRNNHGECCVAVHDKSLRCPHECK
LHCMAKTVPIDKENCCEKVPVCYSNNPLDGGHPT"
19762..21336
/note="Geneid exon scores (in order of location ranges) :
106.68 - GSCU_ID dd_00725"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAS38632.1"
/db_xref="GI:42733688"
/translation="MDNNNLPPQFPFPOPNNYCCDDDNENGNDSTLTIOKPIAPSLPS
SOQSPRTISQIHNKYOOLFKEPKHNNDISKMLCYOLSPQISPTATNPSTPLSN
SLMIPGTFGKLVNHLKTKYIIPYRSRVANNVNSPESFVYDLSLNNALYDXGNNR
STVNNPTNYNNLNLSNNADKOKTITSTATTILANKENVELNNGNNNNLNKNS
EDLIDKSIRKNNKVDGYITRTINRILILLACSSDLSDESLNMCWSQGDENNVERK
SLIDENQFALDQVEVLARPEVDVWKRDYLNHLLEFQSSSELYLQOOSQKRVNV
SSTIQNEKFLSLVYIPFYRYKFEQFSSFNNGNSVARSAPRIGAKIGDFLAK
YTONPVLGNFPEPFSCKGDELATIIDSPVNTSSYPLTFSSSNTPFGSSSIGHL
TIKNSGTSITLTLEGOKRSYQKGNPSSLPLRHRTSSFKHGEFELIIDENVNDIYI
KINREGNTIQSIKYN"
CDS
complement(21961..24357)
/note="Geneid exon scores (in order of location ranges) :
173.48 - GSCU_ID dd_00724"
/codon_start=1
/product="similar to Arabidopsis thaliana (Mouse-ear
cress). Putative PRG81-like negative regulator"
/protein_id="AAS38633.1"
/db_xref="GI:42733689"
/translation="MSKYYSKSQSVSPGKPNQIFVTALPSVSAPSLSLQKPNNGAG
SHRORONSGSSKRPNLPLDOIESTRKSTKNNKRYSKNSKNSNNNIN
SSSIVIKCEBQOQOQOQOQOQOQOQOQOQOQDGLKQEPNLEKNTLYLKDQIS
SPSSSSTSSPSCLSSSSSSTSFQSOSEIAENDQDINNKNCNTNNGNEFEEBEEKNV
EQOSIEETPIESFEETKEVEVPEPEQEIPIPIGEBTLKEYTSPISVISELIV
DMVNDPKSSTSSQSSSSSSSSSPKPIKNNDNNTTTTTTATTATTTTSSKATSS

```

[illegible]

Db 612 TTGTAATTTTGTATTTTAAATATTTATTTATACATATAAACTTTAAACACAC 553

Qy 712 TTTTATTAATGACTATAAGTCAATTTTATATAGACGACGATGTCGAGATTCGCT 771

Db 552 ACACACACACACCAATGTTTTTTTTGTTTTTTTTGTTTTTTTTTATCAACCT 493

Qy 772 TCAAAATCTTCTGATTTTTTAAAGACTAGTTGGCAACCCGTTCTTTCAAGAT 891

Db 492 TTCATATGCTTTGATTTAGTTATTTTATTTTATTTTGAATGTTTTTTTGAAGAA 433

Qy 832 TTGATTTTTCACAAAAATAGTTATTTTCTTTAT 871

Db 432 ATTTTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 393

RESULT 12

LOCUS BX957346/c 149526 bp DNA linear HTG 08-OCT-2004

DEFINITION Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.

ACCESSION BX957346

VERSION BX957346.13 GI:54019944

KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.

SOURCE Danio rerio (zebrafish)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. 1 (bases 1 to 149526)

REFERENCE Beasley,H. Direct Submission Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk On Oct 9, 2004 this sequence version replaced gi:53850295.

COMMENT ----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: zfish-help@sanger.ac.uk ----- Project Information Center project name: ZC117K16 ----- Summary Statistics Assembly program: XGAP4; version 4.5 Chemistry: Dye-terminator; 100% of reads Consensus quality: 149526 bases at least Q40 Consensus quality: 149526 bases at least Q30 Consensus quality: 149526 bases at least Q20 Insert size: 149526; sum-of-contigs Insert size: 149176; 6.5% error; agarose-fp Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality coverage: 12.54x in Q20 bases; agarose-fp -----

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

Source

1. 149526

/organism="Danio rerio"

/mol_type="genomic DNA"

/db_xref="taxon:7955"

/clone_id="CH211-117K16"

/clone_1lb="CHOR1-211"

1. 149526

/note="assembly_fragment:03115"

clone_end:SP6

vector_side:right

clone_end:T7

vector_side:left"

ORIGIN

Query Match 5.1%; Score 70.8; DB 2; Length 149526;

Best Local Similarity 46.8%; Pred. No. 0.00044;

Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

Qy 488 TATCATGTATGTTTAAAGAAATCTTTATTAACACGATCCTTTAAAAACAAGCAT 547

Db 109063 TTTTAAATATGAAACAAAATCATATTTTACTTTGGCCCTTTGTAGCATTAATCAT 109004

Qy 548 ATTTCGAAAGAGCAATATATGTTACA-----GTTTCAAAACATCTAAGGCGCAATTA 603

Db 109003 GTTTTAAAGATTTTAAAAATTTTAAACATGTTTATATGACATATATATGACATTAAGAG 108944

Qy 604 TATCGAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATCTGTTATTTGTTAT 663

Db 108943 CCAAAGTACTATATGACATTTTAAATGTTTATTTTCAAGTTTATTTTATTTTATTT 108884

Qy 664 GTTTTATATACATTTTCTTCTCTTCAATATGAGATGATTTCTCCGATTTTATTAATG 723

Db 108883 AT 108824

Qy 724 ACTATTAAGTCAATTTTATATAGAGCAGCAGCATGCTGATTCGTTCAAAAATCTTT 783

Db 108823 TATATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 108764

Qy 784 CTGATTTTATTAAGACTAGTTTGGCAACCCGTTCTTTCAAGATTTGATTTTTC 843

Db 108763 ATATATCATATTAATAAATCTTTATATGTTATGTTATATATATATATATATATATAT 108704

Qy 844 AAAAAAATTAAGTTATTTTCTTTATTAATAATAGAAAACCTTGAATAATATGAGTTGC 903

Db 108703 AGAAAAACCAAAAGTTAATATTTTATATCTTTATATTTAAATTAATTAATTAATATTTT 108644

Qy 904 CAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACAATCACTGTATATATTTGGCC 963

Db 108643 TATTTATGCTCCTTCAATGATCTTTATATTTTAAATAATATTAATATATATTTTGTG 108584

Qy 964 AGCCCATTAATTTATTTAAACGAAATGAAATCGAACCAATCTGAGCTATTTTC 1023

Db 108583 ACATATATATATTAATTAAGAAAGCAAAAGTAAAAAATGTAATTTTATATATCTGA 108524

Qy 1024 TCTAGATTATGTAATA 1039

Db 108523 GTTTTATATGTTAAAA 108508

RESULT 13

PFMAL4Pl/c 347582 bp DNA linear INV 29-JAN-2003

LOCUS Plasmodium falciiparum MAL4Pl.

DEFINITION AL034557 AL844503

ACCESSION AL034557.8 GI:23498126

VERSION

KEYWORDS Plasmodium falciiparum 3D7

SOURCE Plasmodium falciiparum 3D7

ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE

AUTHORS

1. Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,C., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corton,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Felwell,T., Goble,A., Goodhead,I., Gilliam,R., Hamlin,N., Hance,Z., Harper,D., Hauser,H., Hornby,T., Holroyd,S., Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D., Kerhornou,A., Knights,A., Konfortov,B., Kyes,S., Larke,N., Lawson,D., Leonard,N., Line,A., Maddison,M., McLean,J., Mooney,P., Moule,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabinowitch,E., Rajandream,M.A., Rutter,S., Rutherford,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitehead,S., Woodward,J., Sulston,J.E., Craig,A., Newbold,C. and Barrett,I.B.G.

Sequence of Plasmodium falciiparum chromosomes 1, 3-9 and 13

JOURNAL Nature 419 (6906), 527-531 (2002)

MEDLINE 2255708

PUBMED 12368867
REFERENCE 2 (bases 1 to 347582)
AUTHORS Hamlin, N., Pain, A., Bertman, B., Hall, N., Bowman, S., Churcher, C.,
Harris, B., Harris, D., Lawson, D., Quail, M. and Bartell, B.
TITLE Direct Submision
JOURNAL Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On Oct 3, 2002 this sequence version replaced gi:5731897.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
Location/Qualifiers
1..347582
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4p1"
10384..32958
/note="rep20 Repeats"
join(35153..41725,42757..44124)
/gene="VAR"
/note="synonym: PFD0005w"
join(35153..41725,42757..44124)
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific
surface protein var-2 SMALL:Q26033 (EMBL:L40609) (2664 aa)
fasta scores: E(): 7.8e-149, 43.1% id in 2872 aa"
/product="erythrocyte membrane protein 1 (pFEMP1)"
/protein_id="CAD49094.1"
/db_xref="GI:23498127"
/db_xref="GOA:Q81220"
/db_xref="UniProt/TREMBL:Q81220"
/translation="MVRTGSGGGVGDKIGDIDHQAHLDSIGKKVHDQVNGADGT
GASGAKNIIDLKGDLOKAPINPLIGTDIPCKLVEDYVNNHVNVDGKGRVYCTE
LSGKFPNPFSTLGGQCTNSKMRSGCEGACAPYRHLCHNLSEITTSYADTL
LLEVCMAYEBOGSIINTHTKHSNKSQPSOLCTIARSPADIGDIPVAGKDLFGNT
YESAREKLENTLKEVFGKIHGSLSEAKKQDODGNGYOLREDMTANRETPWAKI
TEVYSGNNYPRATCGDEKNPISLTKQCDKDKAKPIKSGSNVITYPTDYVPOY
LMPFEMWADFCRLRHKLKDAIKCRGNNGSEKCDLNRVCKNTASGKHFFEDP
CDQOYSCAPFVMDINQLEFLKORRKYTKBITSGSGSGSGSKRRKRDATTNVEYE
KKEVLEKGTGKYVNNPLEKLNDEVDCTKNNDIKOGNIDPKNVHSGAKKGDNNK
TYRTKYCEACPMGCAKEVGGWKAKEKNSOTKDIDPDKTTTIELIDGTRKQMVQ
KYKPCNGNGSGEKSAATPNATSRKKKGKCGDMERKICYDENKEXKQSDA.INVCVL
QDGKORKEQKUTSYNAFPWKVYHMLHDSYEMERLNSCTINNAKSONCKNEKNKE
CCGPEKWNQKKEKEWEAKIDHFGKOKDIIETGCDAGVTLAVALLEFLNDELTEKS
EKGLEAEKAEIKHLRQMLEQAGVRLAAGVGPCTEGVAEONTIMDKFLDELEKAE
OCKNCPKPAQOEGPGARSADSPPAGTDHPDAEDDDDDDDDEDEDEDEDEDE
POCKTVNDILSTDRTKQVGDCEKKYGGKNDMGCDLTLVDTRKVCMPRRKQLCL
VYLAHSESTRIETODDLADAFIRTAETPSWYKIKGNADAKOLDNGLIPPEFL
RSMYITGYDRDICLANTDISKTVNDYAKDKIKFRSKGSKSPSGTTTPDDMMQTY
GNDIKWGMICALHGVNTTEKTKIKNDYSYKVNOSONGNSLSDPAKPPFLMMI
EWGEBCFARGKLEONIGKSCNGINPIQYCSNRRHCNACDEBYNAYETKQEKERGQ
TKRFVADANLENADDEYKDTTQGSQGNNDYLKDCDNKSKCMEGVLTLDVSKSK
PGIYAHKYEKCKNCAGKAFVNNVPAPPOPPEPPALPAPATTPGVNVCVQTLF
SDTNKFKACCTLKGPKAPTSMKCIPTGNTNMGAAATSEGSASRHRDIALPSSGS
NGSICVPPRRKLYVTPYTKAAETTKSGSKQBSGKAAGTSESSGSESSSGGTSO
GKSPQGLSTPASTSPNSRDDDLKAFVESAAVETFLMHKYTKADYKDELEKKKO
QESGLVGLDNGSNVSDDEDKOPKLEKGIPEEFPKQMYTTLGDYKDELEKKKO
SDSGTNGSNNNNIIVIEASGDQDEMKIKOAKIDEHINSLOAASVNPQRGOQOON
SILTRTELTKEHAPSIWEGMICALTYKENDEKIKVDNEVEKFPQTTGTGSGYKE
KYEYNTKLDENSDTEAKDTKADABSDNPTFLSHVLABPYRYLMEGGEFCKER
KTLAQIKVDKDSADYKCSGYBECKEDINIGVPADLKPCGGRBEKRYKKYIER
KKIIFEBQSAIVKQTKCKESGGGNGVCGTIVKICDPAOLFRLRGKCKKNDIEG
KLEPNEQSTFTRTKHCDPSCSFKIDCRNGKSGDPTGKCGIITTIKAEIANKMIS
TJDEVNRVSDNTNTEEGDILFYCEBGKGFKGIKRKEWKCRNEGLDVCGLKGNNG
KLDDKQIILIRALIKWLEFLYLEDVNIHGKISDCINNEGNIICRDKONCKNGCYE
ITLKEEMERIKKHVLEKNKEGDNMKS SVRNPLKPEHRBPBNKAIYDKQKLPES
FGLNGDKRSQNGHDAIDCMIKKLEDKITSCLSISTSGQRTTECEHTPLPDDDL
LSETEPNVGKQOPSCFPVKEEKETVDEBKGCEBTSBPRTKIDESSKKKKDSTDSP
BPPPATPAPKEEKKVEPOPOPPTTPOIIVDKTPALVYTLAWSVGIGFAFTYFLL

KKKTSITDLRVINIPKSDYDIPTKLSPNRYIPTYSGKRGKRYTYLEGDSGTSGY
TDHYSITSSSESEEMDINDIYAPRAPKKTILLEVLESPGNNTTSGNNTTASGN
NTTASGNNTTASGNNTTASGNNTTASGNNTTASGNNTTASGNNTTASGNNTTASGN
VNDYSSGSDIPPTQHTLTFDQDEKFPFITSIHDRNLTYGEEVYDMSNGNDLY
NGKNNLYSGQNNVYSGIDPTSDNRGLTSGKSDYSIDILINDTSGNCHIIDYDLK
RKENLPGTNHAKTTIRFAPKPADDDLNHQLLEFHWMLDRHRMKKNKEBEL
KLKEEMNETSHGTHPDSNKTLLATVDSIODMNPPIQOFTMDINVDPTMDNM
EDDIYVDNHDSTVDSNTMDVPSKVQIEMDVNTKLKEXKPIADVMDI"
45450..46601
/gene="PFD0010w"
45450..46601
/gene="PFD0010w"
45450..46601
/note="VAR exon II, could be alternate exon II for
upstream VAR gene PFD0005w"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="synonym: PFD0015c"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="1 probable transmembrane helix predicted by
TMHMM2.0 at aa 328-350
Similar to Plasmodium falciparum rifin pfb0030C
SMALL:Q96112 (EMBL:AE001367) (370 aa) fasta scores: E():
3.4e-65, 53.6% id in 375 aa"
/codon_start=1
/product="RIFIN"
/protein_id="CAD49095.1"
/db_xref="GI:23498128"
/db_xref="UniProt/TREMBL:Q81219"
/translation="MKIHYINILFELPLNLTLYNQRNKHSTNLTNNRSLCECELY
APATYDDPOMKEVWVSKOTQORFHYEDRWMEKMQCQDKKEIOKILKQME
KELMDKFPATLOISDAIPVTCESKSLADRYEGKGLGAVGATVGLISGAV
HWKPKALEAIAKAIAGSTADIAAABAAKAGMEVITAKRGHNGENPPGICDT
ISSGNTYKATFPAVNTIYSKNGTCNMRSSINPACTITIBLSITRGAGTGHP
LYAIRQMTGLAEETBAKAEAKNAKLTAIKERQCTALILEGFSSITNSASTI
AILIIVLMVITVYILIRYRRKKKKQYIKLKE"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="synonym: PFD0020c"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific
surface protein var-3 SMALL:Q26032 (EMBL:L40609) (3006 aa)
fasta scores: E(): 6.2e-103, 43.3% id in 3551 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (pFEMP1)"
/protein_id="CAD49096.1"
/db_xref="GI:23498129"
/db_xref="GOA:Q81218"
/db_xref="UniProt/TREMBL:Q81218"
/translation="MGTSSTPSYVKDYNRSNARVNLNIGIEIYNEKKVNGY
TSOLRGDLSRPFQGLRKARLITVLPAPNSCDLHKFYNINNGYPPARNCDLR
ONRGEANEAYCNSDKIRVTYKSKAGCACAPFRONMCDKALVETINDTNTDITDILG
NLVYTAKEGSIYAKHPHENSSEVCTALASPADIQDIVGKMDPKNEBDAVQGL
SADPKINDNLKEIKESIDYDNDPNYKLRBMWAMRNDQWRALITCYIPYVNVFE
RNPITVFNDEKCGHYEGAPPTNIDYQFLRPMDEMAEEFCRIENHLRIKIKACRN
DKBRLYGQNDYCTKRIEKGSSREKKGCGSKGSDYQFWLKEQNOONPEQTE
DKSILEYVNTKPIGNSNNTKKKYEYFEYELIKQYGSVKNFLOLNKRGCOETIE
EDALIDFTYKGAHAFPSDYQCPPEPCYVEDGCTCTYKTDDEKCRSKTIIOKIISE
TPTEIEVLYSDKCGVITEKIDFCRGNNTNDELQOMKCYNKRSEYNKESIMWLY
QDPKEIMWLEVECFHSAWAKLIDTIRMEHQLKQNLINTVTTDTSKICNCECYA
WIERKKQEWELKEVLANKDETSNHYAKLDVDRFLQYMAFLDOEDCKMDPTE
DLKKRFGSEVSAGTANSDAIEFLDLKONALTGRNNSIKPTCYPPNPDPNGCT
NMNGKLVYRIRALBMMORBYARKOLERKGEINIKADSOGRVIRGGEKXLDNGICN
IDTYSIDNSRNGNGPCTGKNDKGFKGTGTSVGEHEKKTPEVYPPRREHICIS
NLEKLDVSVIRKNNASHSLGADVLLAKYBAKIKELVQNNNSKNVYIDNDKETT
RAMYKSPADIDIRGKDMVQNTDATQLQVLAKEIPKIXDNHDIIGKLOVNGDIT
HKLREBDEWAKNORWAKMCAIENDMONGNPIEDYIPORLWMTWAEVWYCKE
OSRYLGELEKCGSKGCKTEBGDVGKCAKCDYKDEINRWBEWTKIKGYKT
LYKATKRGVYTSNNPDEKVDYFLOGLPKRSKNTPGYVAMPNTLYXSAAGYIHO
ELGKVGCTNTEFCDNKGRYACRHPPKEIEACTICDTRQAOQPIKDKDNCIKT
LIDNSNGTGTGIDCNPKIYGPBSMNCRNSKAKENKACVPPREKCYLLAEGI
FKNKGEDIRETFVXSALETYFAWKRYNDNKKKABEELKSGTIBENFRQWVYTFADY

RDIFGDTIDSHDILIDVSNKAKNKLKREKNGEOKSVIILIDDEKLADMMHGHETME
REBKEKEVEYKCKCKDHESCKNPKTKNGHCYACADYEYISTKKQYNTOKKED
IDKNGKEEYENTKDEKADYLDKCEPGYCDYMEKKNSEYDKNKTYTNDLEK
KCBCKQPPPPAPPOFOSACEIYVDILNGSADYIEKCKGKTYGPEWNCNIOIH
RTHNGKMPRROKLCVYNLQYFKGKTVDLRAFIKCAVEFPFPMHKEKNGE
DLNOLSESGIIPDDFOYMTFGDYDFLFGDISGIGELGKIDSLKPNIGG
KNGDLSRKMNMENGPYIWKMLCSLEKMGDTIKNSNKNYHNVKPNYDNGDL
ETPAKRPOLRLWTEMGDEFCREOKKODLILKCKCPRETTMEGKKCKGADKAYKE
WLOTWKEHYEKOKIKIENDKOSTYNDVTQSPQAYOLNKLKLEKICPSGNTANCEY
KCMKYPSSONNNMNPASLDPTPSDYKDTCECTSOASBNPVSRSBDEGPPPPAP
ROSLARSADNPBRPAPPGPGPPGPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP
OPILARVDOEDEDDE
OEBEPDETEVEYVAPEVKACETVEKLEFNDTKMDEANAKKPEKTYOYKNVLE
PSGDDSTTSVNDGROHRRABGATGKSDAGSICVPRRLRLVGLKJTONASQRT
OGGTSOIVKKTASQPNSHPTLSPSSNPRDGLDAVEIOSAAVEFFMLVGRKMLTK
KPAATLGLGLOIPLAMKALINGVPSGDDNPOKLEGELEPEFKOMPTLDDYDI
LFGKNDIVIGTSGASDKEMKAKEIKETIDKVPNSVSTPPTPATKPSDEKKT
MMEANGEHIMKMIYALTYKNGEKIKYKNEVYKLMDEANAKKPEKTYOYKNVLE
ENGAKPTOPSPSGDNTPTTLNPSIRPVYPRYLBEMGTFPCREKKELEPRKCR
GEYPERKTCGGDHDTENGELOHTMNFALDURDCHQCKRKRMIIDIKFEBEYKOK
DKYQGELOKLGNSNGNCKEIKKHTSASEFLKELKHKQDQNSDDTKSEEDK
NNKIDPNKPLETNPSTYCECTPSNKNVNCNGSGRGRDPCPTPHNEKGSMSVNA
NGNSTEITVEMIDREPLIKNSKILIESGNSDSLFTSRLEKVRDQOMECRYKD
EKTIDCKLKNFNKIDNOYTFPRFLEVWLODFIGSYILKRRKIIIOCKENGGBTC
NENSKNDACVCKWVAOKTTEBNQIKDHTNKEYGNGYDMSHVKYVPEKNEBELKRW

Query Match 5.1%; Score 70.4; DB 3; Length 347582;

Best Local Similarity 44.6%; Pred. No. 0.00051;

Matches 275; Conservative 0; Mismatches 341; Indels 0; Gaps 0;

QY 444 CAAGATTAAAAAATACCAAGTACTATGACCTCGAAGCTATCATGTATGTTA 503

DB 335664 CAAAAGAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAATGTAAG 335605

QY 504 AAGAACATCTATTTAAACACGATCTCTTAAAAAACAAGCATATTTCGAAAGACAA 563

DB 335604 AATATTATCTATGATATTAAGATATTGGAATATTATTTTATGAAAAATATATAT 335545

QY 564 ATTATGTTACAGTTTACAAACATCTAGAGCGACAAATTATGAAAGGTAAAGCTATGA 623

DB 335544 ATATATGTAATAATTTTTTTTATATATATAAATATATATATATATATTTTTTTT 335485

QY 624 CGTTACAGATTTTCTTTTCAATCTGTATTTTGTATTGTTTTATATACATTTCTT 683

DB 335484 TTTTTTTTTTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 335425

QY 684 CTCTTACATAGAGTATTTTCTTCGATTTTAAATGACTATTAAGCATTTTATA 743

DB 335424 TTTTTCAGTATTAAGATTTTAAACGAAATATGTAGAAAGAAAGAAATTTGGTATG 335365

QY 744 TAAAGACGCGATGCTGTAGATTTCTGTTCAAAAATCTTCTGATTTTTTAAAGGTAG 803

DB 335364 TAAATTTTAAATTAATAAATAAAGAAATATATTTTTTTTAAATTAATAAATAAAG 335305

QY 804 TTTGGCAACCTGTTTCTTCAAGAATTTTGATTTTCAAAAAAATTAAGTTATTTT 863

DB 335304 TTTTATTTTATTTTATATATATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 335245

QY 864 CTCTTATTAATAAGAAAAACCTTGAATAATAGATGCGACGCTAGCCCTAGAAATGTT 923

DB 335244 ATATTTTTTTTTTTTTTTTGTCTTATATATATATATATATATATATATATATATTTT 335185

QY 924 TTCCCATTAATTAACATCACTGTATTAATTTTGGCAGGCCCTAATAATTTTAA 983

DB 335184 TTTTTTTTTTTTTTTTGAATATATTTTATTTTATTTTCTACACTATATTAATAATA 335125

QY 984 CCGAACTGAATCGACGAAACCAATCTGAGCTATTTCTTATAGTTAGTAAAAAGGA 1043

DB 335124 TAAATTTTATTTTAAAGGTATCATTAATGTTTTTTTTTAAATTAATAAATAAATAATG 335065

QY 1044 GAGAGAGAGAGAA 1059

DB 335064 TATCTATATATAAAAA 335049

RESULT 14 CO422641/c 883 bp DNA linear PAT 28-JAN-2004

LOCUS Sequence 7675 from Patent WO0151628.

DEFINITION CO422641 GI:41374870

VERSION CO422641.1

SOURCE Homo sapiens (human)

KEYWORDS Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS Lillie, J., Xu, Y., Wang, Y., and Steinmann, K.

TITLE Novel genes, compositions, kits, and methods for identification,

JOURNAL assessment, prevention, and therapy of breast cancer

FEATURES Patent: WO 0151628-A 7675 19-JUL-2001;

source Millennium Pharmaceuticals, Inc. (US)

location/Qualifiers 1. 883

ORIGIN /organism="Homo sapiens"

Query Match 5.0%; Score 70.2; DB 6; Length 883;

Best Local Similarity 39.2%; Pred. No. 0.00071;

Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

QY 419 TGTCCACTTTGGCACTACCCAGACAGATTTAAAAAATACCAAGTAACTATCCA 478

DB 873 TGTNNNNNNNACCTTAAACCTTTAAANNTNNNTTAAANNTTAAANNTTAAANNT 814

QY 479 CTGGAAGCATCATGTAATGTTTAAAGAAACATGATTAACACGATTCCTTAA 538

DB 813 TAAAAATTTTNNAAAACTTAATTTTAAANNTAAANNTTAAANNTTAAANNTTAA 754

QY 539 AACACGATATTTTGAAGAGACAAATTTGTTACGTTTACAAACATCTAAAGAGACA 598

DB 753 TTAANNNTTTTAAANTTATTTAAACCAAAATTTTTTTTAAAAAAATTTTTTAA 694

QY 599 AATTATATGAAAGTAACTATGACGTGAGATTTTCTTTTCACTTGTATTTTG 658

DB 693 AATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 634

QY 659 TTAATGTTTATATATCATTTTCTTCTCTTACATAGAGATTTTCTCCGATTTTATA 718

DB 633 TTTTNNNTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 574

QY 719 AATGACTATTAAGTATTTTATTAAGAGACGATGCTGATGTTCTGTTCAAAA 778

DB 573 AAAAAAANNTTTTTTTTTTATANTTTATTAATAAATAATTTATTTTCTGNAANA 514

QY 779 TCTTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAATTTGAT 838

DB 513 AAAAAAATTTTNNAAATTTTNNAAATTTTNNAAATTTTNNAAATTTTNNAAATTT 454

QY 839 TTTTCAAAAAAATAGTTATTTTCTCTTAAATTAATAAGAAACCTTGAAGAAATAGA 898

DB 453 NNANNTTAAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 394

QY 899 GTTGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATTAATTAATTAATTA 958

DB 393 TTTTNNAAAAATATTTTATTAANTTATNNANNAATTTTAAANTTAANNAATTTTAA 334

QY 959 TGGCCAGCCCATTAATTTTAAACGAACTGAAATCGACGAAACCAATCTGAGCT 1018

DB 333 NTTTAATTAATTAATTAATTTTAAACATTTTAAATTTTATTAATTAATTAATTA 274

QY 1019 ATTT 1022


```

XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other:
SQ
Query Match      100.0%; Score 1394; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2,2e-291;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGGTCTCTATGAAAAAGATGATACATGTCATTCATCCGTTTCTTAAGGTC 60
DB 1 CCAATGGTCTCTATGAAAAAGATGATACATGTCATTCATTCGTTTCTTAAGGTC 60
QY 61 CTTCTTCGCTTATTACTGATGAAATCGGGTTTACAAAAAATTCACCGGTGATAT 120
DB 61 CTTCTTCGCTTATTACTGATGAAATCGGGTTTACAAAAAATTCACCGGTGATAT 120
QY 121 CTGCAATGTTCACTTCTCCCACTCGGTGTCACATTTCTTGATATGTCGTTCCCAT 180
DB 121 CTGCAATGTTCACTTCTCCCACTCGGTGTCACATTTCTTGATATGTCGTTCCCAT 180
QY 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCCTTGGATGGCCCA 240
DB 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCCTTGGATGGCCCA 240
QY 241 CGAAGCGTATCGGTCGTGTGTATCCAGGGGATATATGTCGCCACAAATGTCACCTATA 300
DB 241 CGAAGCGTATCGGTCGTGTGTATCCAGGGGATATATGTCGCCACAAATGTCACCTATA 300
QY 301 TTATATCTCTTATGATATTAATTTTGGAAAAATACAAACTTATCTTTGTGTAT 360
DB 301 TTATATCTCTTATGATATTAATTTTGGAAAAATACAAACTTATCTTTGTGTAT 360
QY 361 GGGCTCAGCATGATTTTGGCTTAAGGCCAGAAATGCGAGGACGACCATGTCATGTC 420
DB 361 GGGCTCAGCATGATTTTGGCTTAAGGCCAGAAATGCGAGGACGACCATGTCATGTC 420
QY 421 TCCACTATTGGCACTACCCAGAAACAAGTTTAAAAAATACCAAGTTAACTATCCACT 480
DB 421 TCCACTATTGGCACTACCCAGAAACAAGTTTAAAAAATACCAAGTTAACTATCCACT 480
QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACCATCTCTTAAAAAA 540
DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACCATCTCTTAAAAAA 540
QY 541 CAAGCATATTTCCAAAAGAGACAATTAATGTTACAGTTTCAACATCTTAAGACGACAAA 600
DB 541 CAAGCATATTTCCAAAAGAGACAATTAATGTTACAGTTTCAACATCTTAAGACGACAAA 600
QY 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTCTGTATTTGTT 660
DB 601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTCTGTATTTGTT 660
QY 661 ATTGTTTTTATATACATTTTCTCTTCAATAGATGATTTTCTTCGATTTTATATAA 720
DB 661 ATTGTTTTTATATACATTTTCTCTTCAATAGATGATTTTCTTCGATTTTATATAA 720
QY 721 ATGACTATAAAGTCATTTTATATTAAGAGACGATGCTAGATTCGTTCAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATTAAGAGACGATGCTAGATTCGTTCAAAATC 780
QY 781 TTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTTCAAAAGAAATTTGATTTT 840
DB 781 TTCTGATTTTTTAAAGCTAGTTGGCAACCCCTGTTTCAAAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACACTTGAAGAAATGAGAT 900
DB 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACACTTGAAGAAATGAGAT 900
QY 901 TGCCAGACTAGCCCTAGATGTTTCCCAATAATTAATCAATCACTGATTAATTAATTTG 960
DB 901 TGCCAGACTAGCCCTAGATGTTTCCCAATAATTAATCAATCACTGATTAATTAATTTG 960
QY 961 GCCAGCCCATTAATTAATTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020
DB 961 GCCAGCCCATTAATTAATTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020

```

```

DB 961 GCCAGCCCATTAATTAATTTAAACGAAACTGAATCGACGAAACCAATCTGACCTAT 1020
QY 1021 TTCTTAGATTTGTAAAAAGGAGAGAGAGAAATCACTTTTAAGTCATTTGCC 1080
DB 1021 TTCTTAGATTTGTAAAAAGGAGAGAGAGAAATCACTTTTAAGTCATTTGCC 1080
QY 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGGTGCTACGTCA 1140
DB 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGGTGCTACGTCA 1140
QY 1141 GGTTCGGACGCTCTGCTGTCATCTCATAGGATATCTACATGCTTTGCAACGTTCTGTC 1200
DB 1141 GGTTCGGACGCTCTGCTGTCATCTCATAGGATATCTACATGCTTTGCAACGTTCTGTC 1200
QY 1201 TTGTTTCATTCGTCGAAGCTTTCCTTATTTGTAAACCAAGAGATTAATCTCCCAACAT 1260
DB 1201 TTGTTTCATTCGTCGAAGCTTTCCTTATTTGTAAACCAAGAGATTAATCTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGAAACAGATCTATT 1320
DB 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGAAACAGATCTATT 1320
QY 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGCATCTTTT 1380
DB 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACCAACCATG 1394
DB 1381 CGTCCACCAACCATG 1394

```

```

RESULT 2
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
AC AAH76333;
XX
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO20016097-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI, 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the MS45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 4; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the MS45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (II) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX

```


CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 2; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 2.4e-290;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCATGGTGTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAAGGCTCC 60
 DB 1 CCATGGTGTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAAGGCTCC 60
 QY 61 CTTCCTCTGCTTATTAAGTGAATCGGGGTTACAAAAAAGCTTCCAGGGGTCATGAT 120
 DB 61 CTTCCTCTGCTTATTAAGTGAATCGGGGTTACAAAAAAGCTTCCAGGGGTCATGAT 120
 QY 121 CTCATGTTCCATCTTCCACCTCGGTGACATTTCTGGATGTGGGTGTTCCAT 180
 DB 121 CTCATGTTCCATCTTCCACCTCGGTGACATTTCTGGATGTGGGTGTTCCAT 180
 QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGAGCACCCATCAAGGGCTTTGCGATGAGCCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACCTTTCGGAGCACCCATCAAGGGCTTTGCGATGAGCCCA 240
 QY 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
 DB 241 CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
 QY 301 TTATTTATCTTTAGATATATTTAATTTTGAATAACAATTAATATCTTTGTGTA 360
 DB 301 TTATTTATCTTTAGATATATTTAATTTTGAATAACAATTAATATCTTTGTGTA 360
 QY 361 GGGCTTCAGATATGATTTTGGCTTAAAGGCCCAAAATGCGAGACCAAGCCATCTAGTG 420
 DB 361 GGGCTTCAGATATGATTTTGGCTTAAAGGCCCAAAATGCGAGACCAAGCCATCTAGTG 420
 QY 421 TCCACTATGGGCTCCAGAACAAAGATTAAATAAATAACAAAGTAACTATGCACT 480
 DB 421 TCCACTATGGGCTCCAGAACAAAGATTAAATAAATAACAAAGTAACTATGCACT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACGATCCTCTTAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACGATCCTCTTAAAAA 540
 QY 541 CAAGCATATTTGCAAGAGACAAATTTGTTACAGTTTACAAACATCTAAGAGCGACAA 600
 DB 541 CAAGCATATTTGCAAGAGACAAATTTGTTACAGTTTACAAACATCTAAGAGCGACAA 600
 QY 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTTGTTATTTGTT 660
 DB 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTTGTTATTTGTT 660
 QY 661 ATTGTTTTATATACATTTTCTTCTTCAATAGATGATTTTCTTCCGATTTTAAAA 720
 DB 661 ATTGTTTTATATACATTTTCTTCTTCAATAGATGATTTTCTTCCGATTTTAAAA 720
 QY 721 ATGACTATAAGCATTTTATATAAGAGACGATGCTGATTTCTGTTCAAAAAATC 780
 DB 721 ATGACTATAAGCATTTTATATAAGAGACGATGCTGATTTCTGTTCAAAAAATC 780
 QY 781 TTCTGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTCTGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAAGAAAACTTGAAGAAATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAAGAAAACTTGAAGAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGATGTTTCCCAATAAATTAACAATCACTGTGATATATTTTG 960
 DB 901 TGGCAGACTAGCCCTAGATGTTTCCCAATAAATTAACAATCACTGTGATATATTTTG 960

QY 961 GCCAGCCCATTAATTTATTAACGAAATGAAATCGAGCGAACCATAATCGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTATTAACGAAATGAAATCGAGCGAACCATAATCGAGCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATGTTCCC 1080
 DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATGTTCCC 1080
 QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
 DB 1081 TGAGATGTGCGGTTTGGCAAGATAGCCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
 QY 1141 GGTTCGGAGCTCTGTGTGATCTGACATGGAATCTATCATAGCTTGTCAACGGTTCGC 1200
 DB 1141 GGTTCGGAGCTCTGTGTGATCTGACATGGAATCTATCATAGCTTGTCAACGGTTCGC 1200
 QY 1201 TTGTTTCATCTGTCAGAGCTTGCCTATTTCTGAACCAAGATTAACCTACCCAAACAT 1260
 DB 1201 TTGTTTCATCTGTCAGAGCTTGCCTATTTCTGAACCAAGATTAACCTACCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAAACATCTATT 1320
 DB 1261 CCATCTTACTCATGCACTTCCATGCAACACGCAATATGTTTCTGAAACATCTATT 1320
 QY 1321 AAAGATCAACAAGCTAGGTTCTCCGCTAGCTTCCCTCTGCTCGCGATCTTTT 1380
 DB 1321 AAAGATCAACAAGCTAGGTTCTCCGCTAGCTTCCCTCTGCTCGCGATCTTTT 1380
 QY 1381 CGTCCACCAACCATG 1394
 DB 1381 CGTCCACCAACCATG 1394

RESULT 4
 ID AAH76332
 AAH76332 standard; DNA; 1394 BP.
 AC AAH76332;
 DT 29-OCT-2001 (first entry)
 XX
 XX
 DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN WO20016097-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albreten MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PS mediating fertility in a male plant.
 XX
 PS Claim 4; Page 46; 50bp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (i) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (ii) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -tissue preferred regulatory region from Z. mays
 XX
 SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 5; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 2,4e-290;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 CCAATGCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTCTTAAGGCTCC 60
DB 1 CCAATGCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTCTTAAGGCTCC 60
QY 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTTACAAAAAATCTTCCAGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTTACAAAAAATCTTCCAGGGTGATGAT 120
QY 121 CTGCAATGTTCACTTCTCCCACTCGGGTTGCAATTTCTTGATGTCGGTGGTCCAT 180
DB 121 CTGCAATGTTCACTTCTCCCACTCGGGTTGCAATTTCTTGATGTCGGTGGTCCAT 180
QY 181 CTGACCGAGGCGCCATCAGACACCTTTCGGGACACCATCAAGGGCTTTGAGTGGCCCA 240
DB 181 CTGACCGAGGCGCCATCAGACACCTTTCGGGACACCATCAAGGGCTTTGAGTGGCCCA 240
QY 241 CGAAGCGTATCGGGTGTGATGCCAGGATATATGTCCTCCCAATGTCACCTATTA 300
DB 241 CGAAGCGTATCGGGTGTGATGCCAGGATATATGTCCTCCCAATGTCACCTATTA 300
QY 301 TTATTAATCTTTGATATTAATTAATTTTGGAAAAATTAACAATTAATCTTTTGCTGA 360
DB 301 TTATTAATCTTTGATATTAATTAATTTTGGAAAAATTAACAATTAATCTTTTGCTGA 360
QY 361 GGGCTCAGCATGATTTTGGCTTAGGGCCAGAAATGAGAGACGACCATGCTAGTG 420
DB 361 GGGCTCAGCATGATTTTGGCTTAGGGCCAGAAATGAGAGACGACCATGCTAGTG 420
QY 421 TCCATTAATGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAATCACT 480
DB 421 TCCATTAATGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAATCACT 480
QY 481 CGAAAGCTATCATGATATGTTTAAAGAACATCTATTAACCAAGTCTCTTAATAAAA 540
DB 481 CGAAAGCTATCATGATATGTTTAAAGAACATCTATTAACCAAGTCTCTTAATAAAA 540
QY 541 CAAGCATATTTGAAAAAGACAAATATGTTACAGTTTCAACATCTTAAGGCGCAAA 600
DB 541 CAAGCATATTTGAAAAAGACAAATATGTTACAGTTTCAACATCTTAAGGCGCAAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTGCAATTTTCTTTTCAATCTTGTATTTTGT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTGCAATTTTCTTTTCAATCTTGTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTTACATAGAGTGATTTTCTTCGATTTTAA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTTACATAGAGTGATTTTCTTCGATTTTAA 720
QY 721 ATGACTATAAGCTATTTTATATAGAGGACGATGCTATATCTGTTCAAAAATC 780
DB 721 ATGACTATAAGCTATTTTATATAGAGGACGATGCTATATCTGTTCAAAAATC 780
QY 781 TTCTGATTTTTTTAAGAGCTATTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
DB 781 TTCTGATTTTTTTAAGAGCTATTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTTCTTTTATTAATAAGAAACACTTGAATAATAGAGT 900
DB 841 TTCAAAAAAATTAAGTTATTTTCTTTTATTAATAAGAAACACTTGAATAATAGAGT 900
QY 901 TGCCAGACTAGCCCTGAATGTTTTCGCAATAAATTAATCAATGCTGTATATTTTGG 960

```

```

DB 901 TGCCAGACTAGCCCTGAATGTTTTCGCAATAAATTAATCAATGCTGTATATTTTGG 960
QY 961 GCCAGCCCATTAATATTTTAAACCGAACTGAATCGAGCAACCAATCTGACTAT 1020
DB 961 GCCAGCCCATTAATATTTTAAACCGAACTGAATCGAGCAACCAATCTGACTAT 1020
QY 1021 TTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCACTTTAAAGTCAATGCTCC 1080
DB 1021 TTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCACTTTAAAGTCAATGCTCC 1080
QY 1081 TGAGATGTGGGTTTGGCAAGCTAGCAACCGTAATCATAGCTCATAGGTCCTAGCTCA 1140
DB 1081 TGAGATGTGGGTTTGGCAAGCTAGCAACCGTAATCATAGCTCATAGGTCCTAGCTCA 1140
QY 1141 GGTTGGGAGCTCTCGTTCATCTCATATGSCATATCATAGCTTGTCAACGGTTCGTC 1200
DB 1141 GGTTGGGAGCTCTCGTTCATCTCATATGSCATATCATAGCTTGTCAACGGTTCGTC 1200
QY 1201 TTGTTTCATGTCGCAAGCCTTGGCTTATTTGAAACCAAGAGATACCTACTCCAAACAT 1260
DB 1201 TTGTTTCATGTCGCAAGCCTTGGCTTATTTGAAACCAAGAGATACCTACTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAATATGTTTCTGSAACAGATCTATT 1320
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACAGCAATATGTTTCTGSAACAGATCTATT 1320
QY 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCTCTCTCTGCTCCGATCTTTT 1380
DB 1321 AAAGATCAACAAGCTAGGCTTCTCCGCTAGCTTCTCTCTCTGCTCCGATCTTTT 1380
QY 1381 CGTCCACGACCATG 1394
DB 1381 CGTCCACGACCATG 1394

```

RESULT 5

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

XX AAH76340;

XX 29-OCT-2001 (first entry)

XX Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; promoter; ds.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albrethen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (i)
 CC comprising nucleotide sequences essential for initiating transcription of

CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment

XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 17.5%; Score 243.6; DB 5; Length 255;
Best Local Similarity 98.4%; Pred. No. 6e-43;
Matches 246; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1145 CGGCACTCTCGTATCTACATGACATACATGCTTGTCAACGGTGTCTTGT 1204
DB 5 CCGCGATCCCGTATCTCAATGCGATCACTCACTGCTTGTCAACGGTGTCTTGT 64

QY 1205 TCCATGCTCCAGCTTCCCTATTTCTGAACCAAGAGATACCTACTCCAAACATCCAT 1264
DB 65 TCCATGCTCCAGCTTCCCTATTTCTGAACCAAGAGATACCTACTCCAAACATCCAT 124

QY 1265 CTTACTCATGCACTTCCATGCAACAGCAGCATATGTTCTCGAACAATCTATTAAAG 1324
DB 125 CTTACTCATGCACTTCCATGCAACAGCAGCATATGTTCTCGAACAATCTATTAAAG 184

QY 1325 ATCACAAGCTAGCGTCTCCCGTAGCTTCCCTCTCTCTGCGCATCTTTTCGTC 1384
DB 185 ATCACAAGCTAGCGTCTCCCGTAGCTTCCCTCTCTCTGCGCATCTTTTCGTC 244

QY 1385 CACCACCATG 1394
DB 245 CACCACCATG 254

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.

AC AAH76334;
XX
DT 29-OCT-2001 (first entry)

DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.

OS Zea mays.
XX
PN WO200160997-A2.

PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.

PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.

PI Albreten MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.

PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.

PS Claim 5; Page 47; 50pp; English.

CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a

CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence

XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match 10.5%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 6.6e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGTGCATCTCAGATGGCATACTACATGCTTGTCAACGGTGTCTTGTCCATCCGC 1213
DB 1 CGTGCATCTCAGATGGCATACTACATGCTTGTCAACGGTGTCTTGTCCATCCGC 60

QY 1214 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCAAACATCCATCTACTCAT 1273
DB 61 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCAAACATCCATCTACTCAT 120

QY 1274 GCACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 1311
DB 121 GCACTTCCATGCAACAGCAGCATATGTTCTCTGAAC 158

RESULT 7
AAL15210/c
ID AAL15210 standard; cDNA; 883 BP.

AC AAL15210;
XX
DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 7667.
XX
KM Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.
XX
PN WO200151628-A2.

PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.

PR 14-JAN-2000; 2000US-0176077P.
XX
PR 14-MAR-2000; 2000US-0189167P.
XX
PR 24-MAR-2000; 2000US-0192099P.

PR 29-MAR-2000; 2000US-0193480P.
XX
PR 15-MAY-2000; 2000US-0205230P.
XX
PR 09-JUN-2000; 2000US-0211315P.
XX
PR 25-JUL-2000; 2000US-0220534P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 1378; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing, treating and encoded
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
Query Match 5.0%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 2.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCAGCTAGCCAGACAAAGATTTAAAAAATTAACCAAGTAATATCCCA 478
DB 873 TGTNNNNNNNACCCTTAAACCTTTAAANATNTNNNANTTNNANNNAAAAAATTTTAAANN 814
QY 479 CTCGAAAGCTATCATGTATGTATTAAAGAAACATCTATTAAAAACGATCTCTTAAAA 538
DB 813 TAAAAATTTTAAAAAATCTAATNTTANNTANNTANNTATTTTNTTAAAAAANNNAAAAA 754
QY 539 AACAGACATATTTTCGAAAGAGACAAATTATGTACGTTTACAAACATCTTAAGAGCGACA 598
DB 753 TTAANNNTTTTAAANTTATTTAAACCAAAATTTTAAAAAATTTTAAAAANTT 694
QY 599 AATTATATCGAAAGTATGAGTATGACGTTTTCAGATTTTCTTTTCACTCTGTATTTTG 658
DB 693 AATTAATTTNATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA 634
QY 659 TTAATGTTTATATACATTTCTCTCTACAAATAGAGATTTTCTTCGATTTTATA 718
DB 633 TTTNTTTNNATTAATAAATTTTAAATTAATAAANNTTTTNTTATNNAATATTTTA 574
QY 719 AAATGACTATTAAGCTATTTTATATTAAGAGACGAGTGTGATGATTCGTTCAAAA 778
DB 573 AAAAAAANNTTTTATATNTTATTAATAAATAATTTTATNTTNTTTTCTNAAAAAN 514
QY 779 TCTTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTTTCTTCAAGATTTTGATT 838
DB 513 AAAAAAATTTTNNAAATNTTTTAAAAACCTTNAATNNAANANAATTTTNTNTT 454
QY 839 TTTTCAAAAAATTTAGTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGA 898
DB 453 NNNANNTTAAAAATTTATTTTNTTANTTCTATTAATTAATAAANNTTAAAAATAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTGTATATAT 958
DB 393 TTTNAAAAAATTAATTTATATNTTATNNAANANAATTTTAANNTNANNNAAAAATTTTA 334
QY 959 TGGCCAGCCCAATTAATTTATTAACGAAATCGAAATCGAGGAAACCAATCTGAGCT 1018
DB 333 NTTTAAATTAATTAATTAATTTAAACNTCATTTNTTATTTTAAATTAATAAATAATTT 274
QY 1019 ATTT 1022
DB 273 NTTT 270
RESULT 8
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
XX ACN85231;
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast Cancer related marker, seq id 6381.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX Homo sapiens.
XX OS
XX PN US2003099974-A1.
XX
XX PD 29-MAY-2003.
XX
XX PF 18-JUL-2002; 2002US-00198846.
XX
XX PR 18-JUL-2001; 2001US-0306220P.

XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2003-787014/74.
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 6381; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (SI). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (II) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?docID=2003099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
Query Match 5.0%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 2.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTTGGCAGCTAGCCAGACAAAGATTTAAAAAATTAACCAAGTAATATCCCA 478
DB 933 TGTNNNNNNNACCCTTAAACCTTTAAANATNTNNNANTTAAANNNAAAAAATTTTAAANN 874
QY 479 CTCGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCGATCTCTTAAAA 538
DB 873 TAAAAATTTTAAAAAATCTAATNTTANNTANNTAATTTTNTTAAAAAANNNAAAAA 814
QY 539 AACAGACATATTTTCGAAAGAGACAAATTTATGTACGTTTACAAACATCTAAGAGCGACA 598
DB 813 TTAANNTTTNTTANNTTATTAACCAAAATTTTAAAAAATTTTAAAAAATTTTAAAAANTT 754
QY 599 AATTATATCGAAAGTATGAGCTAGCGTTCAAGATTTTCTTTTCACTCTGTATTTTG 658
DB 753 AATTAATTTNATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAA 694
QY 659 TTAATGTTTATATACATTTCTCTCTACAAATAGAGTATTTCTTCGATTTTATA 718
DB 693 TTTNTTTNNATTAATAAATTTTAAAAAATTTTAAAAAANNTTTTNTTANNAATATTTTA 634
QY 719 AAATGACTATTAAGCTATTTTATATTAAGAGACGAGTGTGATGATTCGTTCAAAA 778
DB 633 AAAAAAANNTTTTATNTTANNTTANNTAATAAATTTTATNTTNTTTTCTNAAAAANA 574
QY 779 TCTTTCTGATTTTATTAAGCTAGTTTGGCAACCCGTTTCTTCAAGATTTTGATT 838
DB 573 AAAAAAATTTTNNAAATNTTTTAAAAACCTTNAATNNAANANAATTTTNTNTT 514
QY 839 TTTTCAAAAAATTTAGTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGA 898
DB 513 NNNANNTTAAAAATTTATTTTNTTCTATTAATTTAAANNNANNTAATAATAT 454
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAACATCACTGTGTATATAT 958
DB 453 TTTNAAAAAATTAATTTTANNTTANNTANNNAAAAAATTTTAAAAAATTTTAAAAATTTTA 394
QY 959 TGGCCAGCCCAATTAATTTTAAACGAAATCGAAATCGAGGAAACCAATCTGAGCT 1018
DB 393 NTTTAAATTAATTAATTAATTTAAACNTCATTTNTTATTTTAAATTAATAAATAATTT 334
QY 1019 ATTT 1022

Db 333 NTTT 330

RESULT 9

ID ADR04296 standard; DNA; 13400 BP.

AC ADR04296;

DT 04-NOV-2004 (first entry)

DE Corn FT homologue nucleotide sequence SEQ ID NO:63.

KW flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant; floral development; plant sterility; plant fertility; flowering time; plant growth rate; inflorescence architecture; tissue culture morphology; cell division; FT homologue; gene; ds.

XX Zea mays.

XX MO2004067723-A2.

XX 12-AUG-2004.

XX 29-JAN-2004; 2004MO-US002422.

XX 30-JAN-2003; 2003US-00343477.

XX (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.PI Danilevskaya O, Hermon P, Bruggemann E, Shitproun D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;

XX WPI; 2004-580996/56.

PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologs, useful for floral development, e.g. engineering plant flowering
PT time.

XX Claim 6; SEQ ID NO 63; 109pp; English.

CC The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.

SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;

Query Match 5.0%; Score 69.8; DB 13; Length 13400;

Best Local Similarity 73.6%; Pred. No. 6.1e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 811 ACCGTTCTCTTCAAGATTTTGAATTTTTCACAAAAAATAGTATTTTCTCTTGA 870

Db 8559 ACCATTTTATTTTCAAGAGGTTTATTTATCAAGAAAAATAGTCAATTTCTCTTGG 8618

QY 871 TAAATAGAAAACACTAGAAAATAGAGTCCGACACAGCCATGAAATGTTTCCCAA 930

Db 8619 AAAAAATAAAAATCAATGAAAAAATGGGTGTCAAACTAGTCTTATTTAGTTTTCAT 8678

QY 931 T 931

Db 8679 T 8679

RESULT 10

ID AA210551
XX AA210551 standard; DNA; 2657 BP.

AC AA210551;

XX 16-NOV-1999 (first entry)

DE DNA sequence of the P-Zip promoter of maize.

KW P-Zip promoter; male sterile plant; glyphosate tolerance; glyphosate;
KW male reproductive tissue; hybrid seed production; crop outcrossing;
KW flower life; ss.

XX Zea mays.

XX MO9946396-A2.

XX 16-SEP-1999.

XX 09-MAR-1999; 99WO-US005126.

XX 09-MAR-1998; 98US-0077277P.

XX (MONS) MONSANTO CO.

XX Brown SM, Fromm ME;

XX WPI; 1999-551420/46.

PT Production of male sterile plants using a gene encoding glyphosate
PT tolerance, used for, e.g. production of hybrid seed.

XX Disclosure; Fig 1A-B; 54pp; English.

CC The present sequence represents the P-Zip promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC production of a second RNA sequence in a male reproductive tissue.
CC Expression of the second RNA sequence in a male reproductive tissue
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The

CC methode can be used with plants such as corn, wheat, rice, canola, oat,
CC barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
CC sunflower, soybean, tomato, cucumber and squash

XX Sequence 2657 BP; 686 A; 611 C; 677 G; 683 T; 0 U; 0 Other;

Query Match 4.8%; Score 67.2; DB 2; Length 2657;
Best Local Similarity 74.2%; Pred. No. 0.00015;
Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

QY 787 ATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATTTGATTTTCAAA 846
Db ATCTTTTAAAGGCTAGTTGGCAACCATTTTTCAGAGGATTTTCATTTTCCGAAG 1941

QY 847 AAAAATTAGTTTCTTTATATAAATAGAAAACATTAGAAAATAGAGTTGCCAG 906
Db 1942 GGAATATTAGTCAATTTTCCCTTGGAAATATGCAATCCATGGAAATGTGTCCCA 2001

QY 907 ACTAGCCCTAGA 918
Db 2002 ACTAGCCCTAAA 2013

RESULT 11
AAK58751/c
ID AAK58751 standard; cDNA; 6027 BP.

XX AAK58751;
AC AAK58751;
XX 16-AUG-1999 (first entry)

DE Maize d111 gene encoding starch synthase enzyme D11.

XX Search synonym; SSI; D11; d111 gene; maize; transgenic plant; ss.

XX Zea mays.

XX Location/Qualifiers

FT Key

FT misc_feature

FT CDS

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

FT misc_feature

PF 12-NOV-1998; 98WO-US024225.

XX 12-NOV-1997; 97US-00968542.

XX (IOWA) UNIV IOWA STATE RES FOUND INC.

XX Myers AM, James MG;

XX WPI; 1999-327406/27.

XX P-PSDB; AAY06199.

XX Nucleic acid encoding starch synthase enzyme of maize.

PS Claim 1; Page 104-107; 138pp; English.

CC This is the nucleotide sequence of the maize gene d111 (d11). To
CC illustrate the role of the d11 locus in starch biosynthesis, a transposon
CC tagging strategy was used to isolate the gene and describe its
CC polypeptide product. The invention reports tagging of the d11 locus with
CC Mo transposon, cloning and characterization of a portion of the gene, and
CC isolation of a near full-length cDNA (the present sequence). The amino acid
CC sequence (see AAY06199) deduced from this cDNA indicates that d11 codes
CC for a 186 kDa protein extremely similar to potato tuber starch synthase
CC SSI1. Its expression pattern indicates that d11 codes for SSI1 of maize
CC endosperm. The d11 product contains unique sequence features in its N-
CC terminus that may mediate direct interactions with other starch
CC biosynthetic enzymes. Mutations within the maize SSI1 gene affect
CC multiple aspects of starch biosynthesis by disrupting an enzyme complex
CC containing starch synthase(s), starch branching enzyme(s) and possibly
CC starch debranching enzyme(s). The isolated cDNA can be used to provide an
CC enzyme with which to regulate the production of starch, and with which to
CC produce altered or novel forms of starch, e.g. in transgenic plants.
CC Expression of d11 in bacteria and yeasts also modifies glycogen
CC production. Claimed expression vectors comprise the cDNA or fragments of
CC it that code for functional portions of d11

XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

Query Match 4.8%; Score 66.4; DB 2; Length 6027;

Best Local Similarity 77.5%; Pred. No. 0.00027;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAATTTTGATTTTCAAAAAA 651

Db 5606 TCTACGGGCTAGTTGGCAACCCCATTTTTCAGAGGATTTTCCAGAAAAA 5548

QY 852 TTAGTTTATTTCTCTTTAATAATAGAAAACATTAGAAAATAGAGTTGCCAGACTAG 911

Db 5547 TTAGTTTATTTCTCTTTAATAATAGAAAACATTAGAAAATAGAGTTGCCAGACTAG 5488

RESULT 12

ABX09935/c
ID ABX09935 standard; DNA; 6027 BP.

XX ABX09935;

DT 17-FEB-2003 (first entry)

DE DNA encoding maize starch synthase III (D11).

XX Starch; starch synthase; glucan association domain; GLASS; linker domain;

XX LINKR; glucosyl transferase domain; GUTR; C-terminal end; CTEND;

XX granule bound starch synthase; GBSS; morphology; retrogradation;

XX waterbinding; swelling potential; gene; ds.

XX Zea mays.

XX WO200279410-A2.

XX 10-OCT-2002.

XX 29-MAR-2002; 2002WO-US009574.


```

XX PR 30-MAR-2001; 2001US-0279720P.
XX XX
XX PA (BADI ) BASF PLANT SCI GMBH.
XX PI Commuri P, Keeling PL, Ramirez N, Mckean A, Gao Z, Guan H;
XX DR WPI; 2003-040678/03.
XX XX
XX PT New DNA encoding fusion protein consisting of 4 different functional
XX PT domains selected from glucan association domain, linker domain, glucosyl
XX PT transferase domain, and C-terminal end, useful for producing modified
XX PT starches.
XX PS
XX XX Claim 33; Page 225-227; 265pp; English.
XX XX
XX CC The invention describes an isolated DNA molecule encoding a fusion
XX CC protein consisting of 4 different functional domains selected from glucan
XX CC association domain (GLASS), linker domain (LINKR), glucosyl transferase
XX CC domain (GLYTR), and C-terminal end (CTEND) which are operably linked to
XX CC one another. The DNA molecule is useful for expressing in plants
XX CC polypeptides including starch synthase enzymes as fusion proteins with
XX CC improved affinity to starch and modified catalytic capabilities and to
XX CC the in vivo and in vitro synthesis of glucan chains of modified lengths
XX CC as compared to plants producing native starch or starch produced with
XX CC native starch synthases. Expression of the starch synthase fusion
XX CC protein along with granule bound starch synthase (GBSS) will lead to a
XX CC modified starch having an altered or improved morphology, retrogradation,
XX CC waterbinding, or swelling potential of the granules, gel strength,
XX CC adhesiveness, cohesiveness, hardness, elasticity, increased or decreased
XX CC granule size, degree of branching, crystallinity, degree of cross-
XX CC linking, and increased or decreased glucan chain lengths. This sequence
XX CC encodes a starch synthase used in the invention
XX SQ
XX SQ Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
XX
XX Query Match 4.8%; Score 66.4; DB 8; Length 6027;
XX Best Local Similarity 77.5%; Pred. No. 0.00027;
XX Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
XX QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAATTGATTTTCAAAAAA 851
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTCCAGAAAAA 5548
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 852 TTAGTTTATTTCTCTTTATTAATAAGAAAAACATTAGAAAAATGAGTGCAGACTAG 911
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 5547 TTAGTTTATTTCTCTTTATTAATAAGAAAAATCTTTGAAAAAATGAGTTCACTACTAG 5488
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 13
XX ADK12106/c
XX ID ADK12106 standard; cDNA; 6027 BP.
XX XX
XX AC ADK12106;
XX XX
XX DT 03-JUN-2004 (first entry)
XX XX
XX DE cDNA encoding maize starch synthase, DUL.
XX XX
XX KW Maize; dull 1; DUL; starch synthase;
XX KW alpha-1,4-glucosyltransferase catalytic activity; starch production;
XX KW plant; gene; ss.
XX XX
XX OS Zea mays.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 120..514
XX FT /*tag= a
XX FT /product= "DUL"
XX PN US2004049810-A1.
XX XX
XX PD 11-MAR-2004.

```

```

XX XX
XX PF 05-AUG-2003; 2003US-00634262.
XX XX
XX PR 12-NOV-1997; 97US-00968467.
XX PR 12-MAY-2000; 2000US-00554467.
XX XX
XX PA (MYER/) MYERS A M.
XX PA (JAME/) JAMES M G.
XX XX
XX PI Myers AM, James MG;
XX XX
XX DR WPI; 2004-238526/22.
XX DR P-PSDB; ADK12117.
XX XX
XX PT New nucleic acid designated dull1, encoding a starch synthase, useful in
XX PT starch production.
XX PS
XX XX Claim 1; SEQ ID NO 1; 58pp; English.
XX XX
XX CC The present invention relates to the isolation of a maize gene, dull 1
XX CC (DUL), and the polypeptide it encodes. The DUL polypeptide has starch
XX CC synthase activity, and comprises an N-terminal arm region, a C-terminal
XX CC catalytic region, and a region of about 900 amino acids terminating with
XX CC the catalytic region. The C-terminal catalytic region has a catalytic
XX CC domain comprising alpha-1,4-glucosyltransferase catalytic activity. The
XX CC dul polynucleotide sequence is useful in producing starch e.g. from a
XX CC transgenic plant or transfected cell. The present sequence encodes maize
XX CC DUL.
XX SQ
XX SQ Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
XX
XX Query Match 4.8%; Score 66.4; DB 12; Length 6027;
XX Best Local Similarity 77.5%; Pred. No. 0.00027;
XX Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
XX QY 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAATTGATTTTCAAAAAA 851
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTCCAGAAAAA 5548
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX QY 852 TTAGTTTATTTCTCTTTATTAATAAGAAAAACATTAGAAAAATGAGTGCAGACTAG 911
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX DB 5547 TTAGTTTATTTCTCTTTATTAATAAGAAAAATCTTTGAAAAAATGAGTTCACTACTAG 5488
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX RESULT 14
XX ABX35844
XX ID ABX35844 standard; cDNA; 439 BP.
XX XX
XX AC ABX35844;
XX XX
XX DT 20-FEB-2003 (first entry)
XX XX
XX DE Bovine EST associated with lactation/muscle/fat deposition #1009.
XX XX
XX KW Bovine; ss; EST; expressed sequence tag; lactation; IMF;
XX KW muscle deposition; fat deposition; genome mapping; gene identification;
XX KW gene analysis; cattle breeding.
XX XX
XX OS Bos Taurus.
XX XX
XX PN US2002137139-A1.
XX XX
XX PD 26-SEP-2002.
XX XX
XX PF 24-SEP-2001; 2001US-00960352.
XX XX
XX PR 12-JAN-1999; 99US-0115707P.
XX PR 11-JAN-2000; 2000US-00480902.
XX XX
XX PA (BYAT/) BYATT J C.
XX PA (MATH/) MATHIALAGAN N.
XX PA (TAON/) TAO N.
XX PA (WARR/) WARREN W C.

```


XX	Byate JC, Machialagan N, Tao N, Warren WC;
PI	WPI; 2003-110599/10.
DR	
XX	New nucleic acid associated with lactation, and muscle and fat
PT	deposition, useful for genome mapping, gene identification and analysis,
PT	cattle breeding, or for genetically improving cattle.
PS	Claim 2; SEQ ID NO 1009; 245bp; English.
XX	
CC	The invention relates to a purified nucleic acid molecule associated with
CC	lactation or muscle and fat deposition (designated LMPD), derived from
CC	cattle, and the LMPD nucleic acid can specifically hybridise to a second
CC	nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC	appearing as ABX34836-ABX49947, or complements of them. Also included are
CC	(1) a transformed cell having a nucleic acid comprising an LMPD nucleic
CC	acid linked to a promoter and a 3' non-translated sequence that
CC	functions in the cell to cause termination of transcription and addition
CC	of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC	(2) determining a level or pattern of a molecule in a bovine cell or
CC	tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC	of the 15112 nucleic acid sequences or its complement or fragment) with a
CC	complementary nucleic acid molecule obtained from the bovine cell or
CC	tissue, where hybridisation between the marker nucleic acid and the
CC	complementary nucleic acid permits the detection of the molecule; and (b)
CC	detecting the level or pattern of the complementary nucleic acid, where
CC	the detection of the complementary nucleic acid is predictive of the
CC	level or pattern of the molecule. The LMPD nucleic acid is used for
CC	determining a level or pattern of a molecule in a bovine cell or tissue.
CC	It is useful for genome mapping, gene identification and analysis, cattle
CC	breeding, preparation of constructs for use in cattle gene expression, or
CC	for genetically improving cattle. The present sequence is one of the
CC	15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
CC	present sequence was not shown in the specification but was obtained in
CC	electronic format from the USPTO web site:
CC	seqdata.neptco.gov/sequence.html?DocID=20020137139
SQ	Sequence 439 BP; 45 A; 51 C; 56 G; 266 T; 0 U; 1 Other;
	Query Match 4.7%; Score 66.2; DB 8; Length 439;
	Best Local Similarity 54.3%; Pred. No. 0.00015;
	Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
OY	626 TTCAGATTTTCTTTTCATCTGTGTAATTTGGTATTGTTTANATACATTTCCTCT 685
DB	182 TTCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTCTT 241
OY	686 CTACAAATAGAGTGATTTTCTTCGCATTTTATAAATGACTPAAGCATTTTATATA 745
DB	242 TTTTATTTTTTTTTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTT 301
OY	746 AGAGCAGCAAGTCGTGAATCTCGTTCAAATACTTTCTGATTTTTTAAAGAGTAGTT 805
DB	302 TGTCGGGGGAGTGATTTTTTTTTTTTTTTTTTTTATTTTCTTTTTTTTTTTTTTT 361
OY	806 TGGCAACCCTGTTTCTTTCMAAGAATTTTGATTTTTCAAAAAATTAAGTTATTTCT 865
DB	362 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTAATCTTATATTTTTTTTTT 421
OY	866 GTTTATA 872
DB	422 TTCTTTA 428
RESULT 15	
AA187279/C	
ID AA187279 standard; cDNA; 346 BP.	
XX AA187279;	
DT 06-NOV-2001 (first entry)	

```

DE Human polynucleotide SEQ ID NO 7339.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccinating; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukemia;
KW nervous system disorders; arthritis; inflammation; sb.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YF, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
XX
XX P-PsDB; AA007348.
XX
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 7339; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX the encoded proteins (AA000010-AA03910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;
SQ
XX
XX Query Match 4.7%; Score 65.6; DB 4; Length 346;
XX Best Local Similarity 52.6%; Pred. No. 0.00019;
XX Matches 143; Conservative % 0; Mismatches 129; Indels 0; Gaps 0;
XX
XX
XX 446 AGATTAAAAAAATACCAAGTACTCAATCCACTCGAAAGCTATCATGTAAATTTAA 505
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 334 ACAATATAGATATATATTTAAAAATTAAGAATTTAAATTTATTCATAAAAATTTTA 275
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 506 GAAACATCTATTTAAACCCAGATCCTCTTAAAAAAACAAGCATTTTTCGAAAGAGACAAT 565
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 274 GACTTATATATATATATATTTTTTTTTTATAGAAAAATATTTACATTTAAACAATAAAA 215
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 566 TATGTTACAGTTTCAACAACCTTAAGAGCACAATTTATATCGAAAGGTAGCTATGACG 625
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 214 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAT 155
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 626 TTTCGATTTTCTTTTTCATCTCTGTATTTTGTATTTGTTTATATATACATTTTCTCT 685
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 154 TATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 95
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 686 CTTACAAATAGAGTATTTCTTCGATTTAT 717
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX Db 94 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 63
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 210.859 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394
Sequence: 1 cccatgctgcctcatgaaaa.....cttttcgcccaccaccatg 1394

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/lna/5A COMB.seq.*
2: /cgn2_6/ptodata/1/lna/5B COMB.seq.*
3: /cgn2_6/ptodata/1/lna/6A COMB.seq.*
4: /cgn2_6/ptodata/1/lna/6B COMB.seq.*
5: /cgn2_6/ptodata/1/lna/PTUS COMB.seq.*
6: /cgn2_6/ptodata/1/lna/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1394	100.0	1394	3	US-08-880-499-2
2	1389.2	99.7	1394	3	US-08-880-499-1
3	66.4	4.8	6027	2	US-08-968-542C-1
4	66.4	4.8	6027	4	US-08-554-467A-1
5	58.6	4.2	2523	2	US-08-410-784A-3
6	57	4.1	19124	2	US-08-487-826B-13
7	55.8	4.0	279	4	US-09-313-294A-5397
8	55.6	4.0	612	4	US-09-902-540-1357
9	55.6	4.0	1394	3	US-08-880-499-1
10	55.6	4.0	1394	3	US-08-880-499-2
11	55.4	4.0	1055	4	US-09-806-708B-23
12	55.2	4.0	2614	4	US-09-004-056-1
13	55	3.9	396	4	US-09-640-173-53
14	55	3.9	396	4	US-09-713-550-53
15	55	3.9	396	4	US-09-825-294-53
16	55	3.9	396	4	US-09-970-566-53
17	54.6	3.9	55886	4	US-09-949-016-15129
18	53.8	3.9	307	4	US-09-313-294A-4743
19	53.8	3.9	6027	4	US-08-968-542C-1
20	53.8	3.9	6027	4	US-08-554-467A-1
21	52.2	3.7	1141	4	US-09-806-708B-22
22	51.6	3.7	134987	4	US-09-949-016-15348
23	51.6	3.7	134987	4	US-09-949-016-15349
24	51.6	3.7	134987	4	US-09-949-016-15350
25	51.6	3.7	134987	4	US-09-949-016-15507
26	51.6	3.7	134987	4	US-09-949-016-15508
27	51.6	3.7	134987	4	US-09-949-016-15509

C 28	51.4	3.7	1141	4	US-09-806-708B-22	Sequence 22, Appl1
C 29	51.4	3.7	50383	4	US-09-949-016-17600	Sequence 17600, A
C 30	51.4	3.7	129415	4	US-09-949-016-16397	Sequence 16997, A
C 31	51	3.7	601	4	US-09-949-016-156535	Sequence 156535, A
C 32	50.6	3.6	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C 33	50.6	3.6	16573	4	US-09-949-016-14876	Sequence 14876, A
C 34	50.6	3.6	18773	4	US-09-949-016-14164	Sequence 14164, A
C 35	50.4	3.6	731	1	US-08-451-405A-2	Sequence 2, Appl1
C 36	50.2	3.6	832	1	US-09-621-976-2813	Sequence 2813, Ap
C 37	50.2	3.6	317366	4	US-09-949-016-16001	Sequence 16001, A
C 38	50	3.6	1392	4	US-09-257-584-1	Sequence 1, Appl1
C 39	49.8	3.6	2435	3	US-09-306-593-1	Sequence 16110, A
C 40	49.8	3.6	231129	4	US-09-949-016-16110	Sequence 11934, A
C 41	49.8	3.6	266283	4	US-09-949-016-11934	Sequence 11934, A
C 42	49.6	3.6	640681	4	US-09-790-988-1	Sequence 1, Appl1
C 43	49.2	3.5	126176	4	US-09-949-016-16137	Sequence 16137, A
C 44	49.2	3.5	126176	4	US-09-949-016-16138	Sequence 16138, A
C 45	47.8	3.4	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-2

Sequence 2, Application US/08880499
Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnaat W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-2

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY      1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCATATCCGTTTCTTAGGGTCC 60
DB      1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCATATCCGTTTCTTAGGGTCC 60
QY      61 CTTCTTCTGCTTATTACTGACTGGAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
DB      61 CTTCTTCTGCTTATTACTGACTGGAATCGGGGTTTACAAAAAATTCCACGGGTGCATGAT 120
QY      121 CTCATGTTTCACATTCCTCCGACCTCGGTGTCACATTTCTTGGATGTCGGTGGTCCCAT 180
DB      121 CTCATGTTTCACATTCCTCCGACCTCGGTGTCACATTTCTTGGATGTCGGTGGTCCCAT 180
QY      181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCCATCAAGGGCCTTTGCGATGGCCCA 240
DB      181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCCATCAAGGGCCTTTGCGATGGCCCA 240
QY      241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
DB      241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
QY      301 TTATTTCTTTAGATATTTATTTTGGAAAAATTAACAATTAATCTTATCTTTGTGTA 360
DB      301 TTATTTCTTTAGATATTTATTTTGGAAAAATTAACAATTAATCTTATCTTTGTGTA 360
QY      361 GGGCTCAGCATAGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGACCATGTCATGTC 420
DB      361 GGGCTCAGCATAGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGACCATGTCATGTC 420
QY      421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCACT 480
DB      421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCACT 480
QY      481 CGAAGGTATCATGTAATGTTTAAAGAAACATCATTTAAAGACAGATCCTCTTAAAAA 540
DB      481 CGAAGGTATCATGTAATGTTTAAAGAAACATCATTTAAAGACAGATCCTCTTAAAAA 540
QY      541 CAAGCATATTTCCAAAAGAGACAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
DB      541 CAAGCATATTTCCAAAAGAGACAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
QY      601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTTCTGTTATTTGTT 660
DB      601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCATTTCTGTTATTTGTT 660
QY      661 ATTGTTTTATATACATTTTCTCTTACATAGAGTGTCTTTCGATCTTAAATTA 720
DB      661 ATTGTTTTATATACATTTTCTCTTACATAGAGTGTCTTTCGATCTTAAATTA 720
QY      721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCGTAGATTCGTTCAAAAATC 780
DB      721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGTCGTAGATTCGTTCAAAAATC 780
QY      781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTCTTTTCAAAAGATTTGATTTT 840
DB      781 TTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTCTTTTCAAAAGATTTGATTTT 840
QY      841 TTCAAAAAAATTAAGTTTATTTTCTTTTAAATAGAAAAACATAGAAAAATGAGT 900
DB      841 TTCAAAAAAATTAAGTTTATTTTCTTTTAAATAGAAAAACATAGAAAAATGAGT 900
QY      901 TGGCAGACTAGCCCTAGAAATGTTTTCCATATAATTAATCAATCATCTGTATATTTTGG 960
DB      901 TGGCAGACTAGCCCTAGAAATGTTTTCCATATAATTAATCAATCATCTGTATATTTTGG 960
QY      961 GCGAGCCCATTAATTTATTTAAACCGAAATCGAATCGAGCAACCAATCTGACCTAT 1020
DB      961 GCGAGCCCATTAATTTATTTAAACCGAAATCGAATCGAAGCAACCAATCTGACCTAT 1020
QY      1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTCCTCC 1080
DB      1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTCCTCC 1080
QY      1081 TGAGATGTGCGGTTTGGCAACGATAGCACCGTATATCATAGTGTATGTCCTAGTCA 1140

```

```

DB      1081 TGAGATGTGCGGTTTGGCAACGATAGCACCGTATATCATAGTGTATGTCCTAGTCA 1140
QY      1141 GGTTCGGCAGCTCTCGTGTATCTCATGATGAGATACATGATCTGTTCAACGGTTCGTC 1200
DB      1141 GGTTCGGCAGCTCTCGTGTATCTCATGATGAGATACATGATCTGTTCAACGGTTCGTC 1200
QY      1201 TTGTTCCATGTCGCAAGCCTTGCATCTGTAACCAAGAGGATPACTACTCCCAACAT 1260
DB      1201 TTGTTCCATGTCGCAAGCCTTGCATCTGTAACCAAGAGGATPACTACTCCCAACAT 1260
QY      1261 CCATCTTACTCATGACACTTCCATGCAACACGACATATGTTTCTGAAACAGATCTATT 1320
DB      1261 CCATCTTACTCATGACACTTCCATGCAACACGACATATGTTTCTGAAACAGATCTATT 1320
QY      1321 AAAAGATCACAAAGCTAGCGTTTCCCGTACGTTCCCTCTCTCTGCGCATCTTTT 1380
DB      1321 AAAAGATCACAAAGCTAGCGTTTCCCGTACGTTCCCTCTCTCTGCGCATCTTTT 1380
QY      1381 CGTCCACACCATG 1394
DB      1381 CGTCCACACCATG 1394

RESULT 2
US-08-880-499-1
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Hufman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-1

Query Match          99.7%; Score 1389.2; DB 3; Length 1394;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```


QY 1 CCATGCTCTCTATGAAAAAGATGATCAATGTCTATATCCGTTTCTTAGGCTCC 60
DB 1 CCAATGGTCTCTATGAAAAAGATGATCAATGTCTATATCCGTTTCTTAGGCTCC 60
QY 61 CTTCTTCTGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 CTTCTTCTGCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 CTCGATGTCACCTTCTCCACCTCGGCTTGCACTTCTTGATGATGATGATGAT 180
DB 121 CTCGATGTCACCTTCTCCACCTCGGCTTGCACTTCTTGATGATGATGATGAT 180
QY 181 CTGACCCAGGCCCCATGACACCTTTGGGACACCCATCAAGGCTTTGGATGGCCCA 240
DB 181 CTGACCCAGGCCCCATGACACCTTTGGGACACCCATCAAGGCTTTGGATGGCCCA 240
QY 241 CGAGACGTATCGGGTCTGGTGTATCGAGGGATATATGTCCCCCAATGCTACCTATA 300
DB 241 CGAGACGTATCGGGTCTGGTGTATCGAGGGATATATGTCCCCCAATGCTACCTATA 300
QY 301 TTATATCTTTAGATATATTTAATTTTGAATAAATAAATAAATAAATAAATAAATA 360
DB 301 TTATATCTTTAGATATATTTAATTTTGAATAAATAAATAAATAAATAAATAAATA 360
QY 361 GGGCTTCAGCATGATTTTGGCTTAGGGCCCAAAATGCGAGACCGCATGCTTAGTG 420
DB 361 GGGCTTCAGCATGATTTTGGCTTAGGGCCCAAAATGCGAGACCGCATGCTTAGTG 420
QY 421 TCCACTTATGGCACTACCCAGAACAGATTTAAATAAATAAATAAATAAATAAATAA 480
DB 421 TCCACTTATGGCACTACCCAGAACAGATTTAAATAAATAAATAAATAAATAAATAA 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACATCTCTTAAATAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACACATCTCTTAAATAA 540
QY 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTAAACATCTTAAAGACGACAA 600
DB 541 CAAGCATATTTCCAAAGAGACAAATTTATGTTACAGTTTAAACATCTTAAAGACGACAA 600
QY 601 TTATATGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 TTATATGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 661 ATTGTTTATATACATTTTCTCTTACAAATGATGATGATGATGATGATGATGAT 720
DB 661 ATTGTTTATATACATTTTCTCTCTTACAAATGATGATGATGATGATGATGAT 720
QY 721 ATGACTATTAAGCATTTTATATATAGAGACGATGCTGATGCTGCTTCAAAAATC 780
DB 721 ATGACTATTAAGCATTTTATATATAGAGACGATGCTGATGCTGCTTCAAAAATC 780
QY 781 TTTCTGATTTTATAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTT 840
DB 781 TTTCTGATTTTATAGAGCTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTT 840
QY 841 TTCAAAAAAATTAATTTTCTCTTAAATAAATAAATAAATAAATAAATAAATAA 900
DB 841 TTCAAAAAAATTAATTTTCTCTTAAATAAATAAATAAATAAATAAATAAATAA 900
QY 901 TGGCAGCTAGCCCTAGATGTTTCCCAATTAATTAATTAATTAATTAATTAAT 960
DB 901 TGGCAGCTAGCCCTAGATGTTTCCCAATTAATTAATTAATTAATTAATTAAT 960
QY 961 GCCAGCCCATTAATTTTAAACCGAACTGAATGAGGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCATTAATTTTAAACCGAACTGAATGAGGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080

QY 1081 TGAATGTGGCTTTGGCAACGATAGCAACCGTAATCATATGATAGTGGCTAGCTCA 1140
DB 1081 TGAATGTGGCTTTGGCAACGATAGCAACCGTAATCATATGATAGTGGCTAGCTCA 1140
QY 1141 GGTTCGGAGCTCTGCTGATCATCTCAATGAGCATATCAATGCTTTCAGACCTTCGTC 1200
DB 1141 GGTTCGGAGCTCTGCTGATCATCTCAATGAGCATATCAATGCTTTCAGACCTTCGTC 1200
QY 1201 TTGTTTCATGTCGAAGCTTGGCTTATTTGAAACGAAGAGATTAATCTCCCAACAT 1260
DB 1201 TTGTTTCATGTCGAAGCTTGGCTTATTTGAAACGAAGAGATTAATCTCCCAACAT 1260
QY 1261 CCATCTTACTCAGCACTTCCATGCAACACGCAATATGTTTCTGAAACAGATCAT 1320
DB 1261 CCATCTTACTCAGCACTTCCATGCAACACGCAATATGTTTCTGAAACAGATCAT 1320
QY 1321 AAGATCACAACAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGCATCTTTT 1380
DB 1321 AAGATCACAACAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACACCATG 1394
DB 1381 CGTCCACACCATG 1394

RESULT 3
US-08-968-542C-1/C
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dulli Codes For A No. 5981728el Starch
; TITLE OF INVENTION: Synthase
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm

IMMEDIATE SOURCE:
LIBRARY: maize endosperm cDNA library in
LIBRARY: (pZf11
CLONE: pMgf10; pMgf6a; pMgf6-2M
US-08-968-542C-1

Query Match 4.8%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTGTGGCAACCTGTTCTTCAAGATTGATTGTTTCAAAAAA 851

DB 5606 TCTACGGGCTAGTGTGGCAACCCCATTT-TTCCAGGGGATTTCATTGTTCCAGAAAA 5548

QY 852 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 911

DB 5547 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 5488

RESULT 4

US-09-554-467A-1/C
Sequence 1, Application US/09554467A

Patent No. 6639125

GENERAL INFORMATION:

APPLICANT: Myers, Alan M.

APPLICANT: James, Martha G.

TITLE OF INVENTION: dult Coding for a No. 6639125el Starch Synthase and Uses

TITLE OF INVENTION: Theoreof

FILE REFERENCE: D6036PCT

CURRENT APPLICATION NUMBER: US/09/554,467A

CURRENT FILING DATE: 2000-05-12

PRIOR APPLICATION NUMBER: PCT/US98/24225

PRIOR FILING DATE: 1998-11-12

PRIOR APPLICATION NUMBER: US 08/062,102

PRIOR FILING DATE: 1997-11-12

NUMBER OF SEQ ID NOS: 37

SEQ ID NO 1

LENGTH: 6027

TYPE: DNA

ORGANISM: maize

FEATURE:

OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the

OTHER INFORMATION: starch synthase enzyme DUL.

US-09-554-467A-1

Query Match 4.8%; Score 66.4; DB 4; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTGTGGCAACCTGTTCTTCAAGATTGATTGTTTCAAAAAA 851

DB 5606 TCTACGGGCTAGTGTGGCAACCCCATTT-TTCCAGGGGATTTCATTGTTCCAGAAAA 5548

QY 852 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 911

DB 5547 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 5488

RESULT 5

US-08-410-784A-3/C
Sequence 3, Application US/08410784A

Patent No. 5912413

GENERAL INFORMATION:

APPLICANT: MYERS, ALAN M.

APPLICANT: JAMES, MARTHA G.

TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING

TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE

TITLE OF INVENTION: SUGARY 1

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP

STREET: Ten Post Office Square

CITY: Boston

Query Match 4.8%; Score 66.4; DB 4; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.2e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTGTGGCAACCTGTTCTTCAAGATTGATTGTTTCAAAAAA 851

DB 5606 TCTACGGGCTAGTGTGGCAACCCCATTT-TTCCAGGGGATTTCATTGTTCCAGAAAA 5548

QY 852 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 911

DB 5547 TTAGTTATTTTCTCTTAAATAAGAAAACTAGAAAAATAGAGTCCAGACTAG 5488

RESULT 6

US-08-487-826B-13/C
Sequence 13, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chlitis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

Query Match 4.2%; Score 58.6; DB 2; Length 2523;
Best Local Similarity 71.4%; Pred. No. 0.00013;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 774 AAAATCTTCTGATTTTAAAGCTAGTGTGGCAACCTGTTCTTCAAGAAATT 833

DB 279 ACATATACATATACATATATATAGGTCTAGTTGCAATCCATTTTCAAAAAG--TTT 222

QY 834 TGATTTTCAAAAAATTAATTTCTCTTAAATAAGAAAACTAGAAAA 893

DB 221 TACATTTTCAAAATTAATTAATTTCTCTTGA-AAAATAGAAATTTCTCAGAAAA 163

QY 894 ATAGAGTTCAGAGTCCAGACTAGCCCTAGAT 920

DB 162 ATAGAGTTCAGAGTCCAGACTAGCTTAAAT 136

RESULT 7

US-08-410-784A-3/C
Sequence 3, Application US/08410784A

Patent No. 5912413

GENERAL INFORMATION:

APPLICANT: MYERS, ALAN M.

APPLICANT: JAMES, MARTHA G.

TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING

TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE

TITLE OF INVENTION: SUGARY 1

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP

STREET: Ten Post Office Square

CITY: Boston

COUNTRY: US

RESULT 9
US-08-880-499-1/c
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-1
Query Match 4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00059;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 761 TAGATTCTGTTCAAAAATCTTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTTTC 820
DB 950 TACACAGTATGTATTTATTTATGGGAAACATCTAGGGCTAGCTGGCAACTCTATTTT 891
QY 821 TTTCAGAGATTTGATTTTTCAAAAAATAGTTATTTCTCTTTATAAATAGAA 880
DB 890 TCTAAGTGTTCATTTTATAAGAGAAATTAACCTAATTTTGTGAAAAATCAAAA 831
QY 881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTCCCATTAATACAA 940
DB 830 TTCTTTGAAAGAAACAGGGTGGCCAAACTAGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761
RESULT 10
US-08-880-499-2/c
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:

; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnaat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-880-499-2
Query Match 4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00059;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 761 TAGATTCTGTTCAAAAATCTTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTTTC 820
DB 950 TACACAGTATGTATTTATTTATGGGAAACATCTAGGGCTAGCTGGCAACTCTATTTT 891
QY 821 TTTCAGAGATTTGATTTTTCAAAAAATAGTTATTTCTCTTTATAAATAGAA 880
DB 890 TCTAAGTGTTCATTTTATAAGAGAAATTAACCTAATTTTGTGAAAAATCAAAA 831
QY 881 AACACTTAGAAAAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTCCCATTAATACAA 940
DB 830 TTCTTTGAAAGAAACAGGGTGGCCAAACTAGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761
RESULT 11
US-09-806-708B-23
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133

	Query Match	3.9%	Score 55;	DB 4;	Length 396;
	Best Local Similarity	43.8%;	Pred. NO.	0.00052;	
	Matches 139;	Conservative 0;	Mismatches 178;	Indels 0;	Gaps 0;
Qy	632 TTTTCTTTTCATCTGTGTAATTGGTATGTTTTATAACACTTTCTCTCTACA	691			
Db					
	11 TTTTCTTTTCATCTGTGTAATTGGTATGTTTTATAACACTTTCTCTCTCTA	70			
Qy	692 ATAGAGTATTTCTTCGCAATTTATAAAGCATATAAGCATTATTAAGAACA	751			
Db					
	71 TTTTCTTTTCATCTGTGTAATTGGTATGTTTTATAACACTTTCTCTCTCTA	130			

Accession	Sequence	Position
QY	TTTTCTTTTCATCTCTGTATATTTGGTATTTTATATATACATTTCTCTCTTAC	691
Db	TT	70
QY	ATAGAGTCATTTCTCCGATTTTATAAAGATCATTAAGTCATTTTATATAGACA	751
Db	TT	130
QY	CGCATGTCGTAGATCTCGTCCAAAAATCTTCATTTTAAAGCGTAGTTGGCA	811
Db	TT	150
QY	CCCGTTTCTTCAAGAATTTGATTTTTTCAAAAAAATAGTTATTTCTCTTTAT	871
Db	TT	250
QY	AAAAATAGAAAACCTTAGAAAAATAGATGCGCAGACTGAGCCCTAGATGTTCCCAAT	931
Db	TT	310
QY	AAATTCANAAAAAGANAAAGAAAAAATAAANANNANCNANNNNNNNNATNTNTNCTTNATA	310
QY	AAATTCACATCATCTGTG	948
Db	NTNTNTNNNNNANNGG	327

Search completed: September 15, 2005, 08:25:12
Job time : 212.859 secs

This Page Blank (uspro)

Sat Sep 17 09:06:46 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 23.7481 Seconds

(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157
1 cgcgcattccacatggcat.....cgcacatattgttcccgaaac 157

Sequence:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB. seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB. seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB. seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB. seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB. seq.*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl1
2	157	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl1
3	30.4	19.4	385136	4 US-09-949-016-16073	Sequence 16073, A
4	29.6	18.9	1185	4 US-09-543-681A-1081	Sequence 1081, Ap
5	29.2	18.6	1830121	4 US-09-543-681A-1081	Sequence 1, Appl1
6	29.2	18.6	1830121	4 US-09-543-681A-1081	Sequence 1, Appl1
7	28.8	18.3	64190	4 US-09-949-016-14712	Sequence 14712, A
8	28.8	18.3	64190	4 US-09-949-016-14712	Sequence 14712, A
9	28	17.8	11032	4 US-09-949-016-15502	Sequence 15502, A
10	28	17.8	11032	4 US-09-949-016-15502	Sequence 15502, A
11	27.8	17.7	601	4 US-09-949-016-17447	Sequence 17447, A
12	27.8	17.7	11490	4 US-09-949-016-191201	Sequence 191201, A
13	27.8	17.7	11490	4 US-09-949-016-16740	Sequence 16740, A
14	27.4	17.5	1413	4 US-09-710-279-1813	Sequence 1813, Ap
15	27.4	17.5	1413	4 US-09-710-279-1813	Sequence 1413, Ap
16	27.4	17.5	1794	4 US-09-270-767-10489	Sequence 10489, A
17	27.4	17.5	242	1 US-08-425-295A-2	Sequence 2, Appl1
18	27.4	17.5	2926	4 US-09-710-279-3747	Sequence 3747, Ap
19	27.4	17.5	3014	4 US-09-710-279-3680	Sequence 3680, Ap
20	27.4	17.5	3599	4 US-09-710-279-4003	Sequence 4003, Ap
21	27.2	17.3	3445	4 US-09-949-016-2556	Sequence 2556, Ap
22	27.2	17.3	3588	4 US-09-566-821-23	Sequence 23, Appl1
23	27.2	17.3	6751	1 US-07-882-925A-5	Sequence 5, Appl1
24	27.2	17.3	6751	1 US-08-184-012C-5	Sequence 12703, A
25	27.2	17.3	8939	4 US-09-949-016-12703	Sequence 12703, A
26	27.2	17.3	9076	4 US-09-949-016-17368	Sequence 17368, A
27	27.2	17.3	48994	4 US-09-949-016-14091	Sequence 14091, A

28	27.2	17.3	60593	4 US-09-949-016-13779	Sequence 13779, A
29	27.2	17.3	62776	4 US-09-949-016-17576	Sequence 17576, A
30	27.2	17.3	100848	4 US-09-596-002-39	Sequence 39, Appl1
31	27.2	17.3	158735	4 US-09-949-016-11989	Sequence 11989, A
32	27.2	17.3	158735	4 US-09-949-016-17130	Sequence 17130, A
33	27.2	17.2	879	3 US-09-134-001C-678	Sequence 678, App
34	27	17.2	44180	4 US-09-949-016-12988	Sequence 12988, A
35	27	17.2	47030	4 US-09-949-016-13037	Sequence 13037, A
36	27	17.2	47030	4 US-09-949-016-13038	Sequence 13038, A
37	27	17.2	47030	4 US-09-949-016-15039	Sequence 15039, A
38	27	17.2	47030	4 US-09-949-016-15040	Sequence 15040, A
39	27	17.2	194537	4 US-09-949-016-12928	Sequence 12928, A
40	27	17.2	201529	4 US-09-949-016-12740	Sequence 12740, A
41	26.8	17.1	265	4 US-09-513-999C-8379	Sequence 8379, Ap
42	26.8	17.1	271	3 US-08-946-026-9	Sequence 9, Appl1
43	26.8	17.1	3114	3 US-08-946-026-12	Sequence 12, Appl1
44	26.8	17.1	156942	4 US-09-949-016-12227	Sequence 12227, A
45	26.8	17.1	156950	4 US-09-949-016-15946	Sequence 15946, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garmaat W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

TITLE OF INVENTION: AND METHOD OF USING SAME

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 157; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5.6e-45;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTATCTACATGGCATACTACATGCTGTTCACCGCTGCTGTTCATGCTGC 60
DB 1155 CGTGTATCTACATGGCATACTACATGCTGTTCACCGCTGCTGTTCATGCTGC 1214
QY 61 AAGCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 120
DB 1215 AAGCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 1274
QY 121 CAACCTTCATGCAAAACGCAACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAAAACGCAACATATGTTCTCGAAC 1311

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garneat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendal, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 157; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5, 6e-45;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTATCTACATGGCATACTACATGCTGTTCACCGCTGCTGTTCATGCTGC 60
DB 1155 CGTGTATCTACATGGCATACTACATGCTGTTCACCGCTGCTGTTCATGCTGC 1214
QY 61 AAGCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 120
DB 1215 AAGCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 1274
QY 121 CAACCTTCATGCAAAACGCAACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAAAACGCAACATATGTTCTCGAAC 1311

DB 1275 CAACCTTCATGCAAAACGCAACATATGTTCTCGAAC 1311

RESULT 3
US-09-949-016-16073/C
Sequence 16073, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 16073
LENGTH: 385136
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(385136)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16073

Query Match 19.4%; Score 30.4; DB 4; Length 385136;
Best Local Similarity 57.3%; Pred. No. 14;
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 55 TCGTCCAGCCCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTA 114
DB 253908 TCGTCCAGCCCTTGGCTATTCTGAACCAAGAGATACCTACTCCCAACATTCATCTTA 253849
QY 115 CTCATGCAACTTCCTCATGCAACGCGACATATGTTT 150
DB 253848 TCCATTTACTCTCCACACACACACATTTT 253813

RESULT 4
US-09-543-681A-1081/C
Sequence 1081, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 1081
LENGTH: 1185
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-1081

Query Match 18.9%; Score 29.6; DB 4; Length 1185;
Best Local Similarity 59.5%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 35 CAACGCTTGTGTTGTTCCATGCTCAAGCCTTGGCTATTCTGAACCAAGAGATATCTTA 94
DB 940 CAACATTTGTTTGTCTCCACCAAGAAAGCCTTAACCCATTTTCTCCAAAGACATATCTTA 881

QY 95 CTCGCAACATCATCTACTCA 118
DB 880 ATCCCTAGATGCTTTTATATA 857

RESULT 5

US-09-557-884-1/c

Sequence 1, Application US/09557884

Patent No. 6506581

GENERAL INFORMATION:

APPLICANT: Fleischmann et al.

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS V6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/476,102

FILING DATE: JUN-5-1995

ATTORNEY/AGENT INFORMATION:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8504

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;

Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 12 ACATGCGATCTACGCTGTTCAACCGTCTGCTGTTCCATCGTCCAAAGCTTGCCCT 71

DB 677275 ACTTTAAATTCAGCGCTGCTGCTGCTTAATTCATTTGATTAAATGCAAGATAGCCCT 677216

QY 72 ATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATGCACTTCAGT 131

DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

QY 132 CAACACGCA 141

DB 677155 CAACCAACCA 677146

RESULT 6

US-09-643-990A-1/c

Sequence 1, Application US/09643990A

Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS V6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-643-990A-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;

Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 12 ACATGCGATCTACGCTGTTTCAACCGTCTGCTGTTCCATGCTCCAAAGCTTGCCCT 71

DB 677275 ACTTTAAATTCAGCGCTGCTGCTGCTTAATTCATTTGATTAAATGCAAGATAGCCCT 677216

QY 72 ATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATGCACTTCAGT 131

DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156

QY 132 CAACACGCA 141

DB 677155 CAACCAACCA 677146

RESULT 7

US-09-949-016-14712/c

Sequence 14712, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14


```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14712
; LENGTH: 64190
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14712
```

```

Query Match      18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```

QY 55 TCGTCCAGCCTTGCTATTCGACCAAGAGATACCTACCTCCCAACATCCATCTTA 114
    |||||
Db 355 TCATCCAAATGTTCCACTACTAATCCATGATGGAAGTACTAATCAAGCAATATT 296
    |||||
QY 115 CTCATGCAACTT 126
    |||||
Db 295 CTCATGTAAACAT 284
```

```

RESULT 8
US-09-949-016-14713/c
; Sequence 14713, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14713
; LENGTH: 64190
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14713
```

```

Query Match      18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```

QY 55 TCGTCCAGCCTTGCTATTCGACCAAGAGATACCTACCTCCCAACATCCATCTTA 114
    |||||
Db 355 TCATCCAAATGTTCCACTACTAATCCATGATGGAAGTACTAATCAAGCAATATT 296
    |||||
QY 115 CTCATGCAACTT 126
    |||||
Db 295 CTCATGTAAACAT 284
```

```

RESULT 9
US-09-949-016-15502
; Sequence 15502, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15502
; LENGTH: 11022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15502
```

```

Query Match      17.8%; Score 28; DB 4; Length 11022;
Best Local Similarity 63.2%; Pred. No. 22;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
```

```

QY 15 TGGCATCTACATGCTTGTTCAACCGTTCGTCCTTCATGTCACAGCCTTGCTATT 74
    |||||
Db 4432 TGGCTTATTTCATGGTGTTTGAAGGTTAATGTTGAGCATCATTTAGCCTTGCAAG 4491
    |||||
QY 75 CTGAACCA 82
    |||||
Db 4492 CTGAATCA 4499
```

```

RESULT 10
US-09-949-016-17447
; Sequence 17447, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17447
; LENGTH: 41815
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17447
```

```

Query Match      17.8%; Score 28; DB 4; Length 41815;
Best Local Similarity 58.3%; Pred. No. 38;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 62 AGCCTTGCTATTCGACCAAGAGATACCTACTCCCAACATCTTACTCATGTC 121
    |||||
Db 33058 AGACATGGAGATTCAACACATCCGATAGCTGCTCTAAGCATCCCTGCTCATTC 33117
    |||||
QY 122 AACTTCATGCAACACGCACATA 145
    |||||
Db 33118 ATGTCATAGAGACATGATGTA 33141
```

```

RESULT 11
US-09-949-016-191201/c
; Sequence 191201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```



```
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 191201
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-191201
```

```
Query Match      17.7% Score 27.8; DB 4; Length 601;
Best Local Similarity 58.0%; Pred. No. 7.9;
Matches 47; Conservative 1; Mismatches 33; Indels 0; Gaps 0;
```

```
QY 57 GTCCAGCCTTGCTTATTCGACCAAGAGATACCTACCCCAACATCATCTTACT 116
DB 322 GTCTGACCTTGATGACACRTGACGCCGACCAACACACACACTCTCTCAT 263
QY 117 CATGCACTTCATGCAACA 137
DB 262 CATGCAATAGATGAAATA 242
```

```
RESULT 12
US-09-949-016-14149/c
Sequence 14149, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14149
LENGTH: 11490
TYPE: DNA
ORGANISM: Human
US-09-949-016-14149
```

```
Query Match      17.7% Score 27.8; DB 4; Length 11490;
Best Local Similarity 52.1%; Pred. No. 26;
Matches 62; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
QY 38 CCGTTGCTTGTTCATCTCCAGCCTTGCTTATTCGAACCAAGAGATACCTACT 97
DB 5642 CAGGGCTCTGCTTCCCTGCTCCCTCCCTCCGACCACTCAGAGAGGAGCCACCC 5583
QY 98 CCAACATCATCTTACTGATGCACTTCATGCAACGACGACATATGTTTCCGAA 156
DB 5582 CCTTGACACCTGATGACATCCGCTTAGTGAGGCGAGGAGATGATGTCCTCAA 5524
```

```
RESULT 13
US-09-949-016-16740
```

```
Sequence 16740, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 16740
LENGTH: 14952
TYPE: DNA
ORGANISM: Human
US-09-949-016-16740
```

```
Query Match      17.7% Score 27.8; DB 4; Length 14952;
Best Local Similarity 59.5%; Pred. No. 29;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY 67 TGCCTATTCGAACCAAGAGATACCTACTCCCAACATCATCTTACTGCACTT 126
DB 70 TGTCAACCGTGAATCCAGATTAATTAATTTAGGATTCCTCTCAGAGTAACTA 129
QY 127 CCATGCAACAGCACATA 145
DB 130 GGTTTACACACACACA 148
```

```
RESULT 14
US-09-710-279-1813/c
Sequence 1813, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMBERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PUS480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1813
LENGTH: 1413
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1813
```

```
Query Match      17.5% Score 27.4; DB 4; Length 1413;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 14 ATGGATATACATGCTGTTTCACCGTTCGCTTGTTCATGTCGCAAGCTTGCTAT 73
DB 1213 ATTGAATCTAGTGTCTGTTTCACAGTTTGTATCTTCATCTGATATTCATCATAC 1154
QY 74 TCTGAACCAAGAGATACCTACTCCCAACAT 106
DB 1153 CAAGATGCAATATATCTGTACTCTCTAT 1121
```

```
RESULT 15
```


US-09-134-001C-1413/C
; Sequence 1413; Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1413
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1413

Query Match 17.5%; Score 27.4; DB 3; Length 1428;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 14 ATGGCATACTACATGCTGTGTTCAACGCTTGCTTGTTCATGCTCCAGCCTTGCTAT 73
DB 1228 ATGAAATCTACGCTCGCTTCACAGATTGTTATCTTCATCTGATAATTCAATCCATAC 1169
QY 74 TCTGAACCAAGAGATACCTACTCCCAACCAAT 106
DB 1168 CAAGAAATCGCAATATATCTTGTAACTCTCTAT 1136

Search completed: September 15, 2005, 08:25:20
Job time : 28.7481 secs


```

Db      1  CAAATGCTCTATGAAAAAGATGAGTAAATGTCATATCCGTTTCTTAAGGTC 60
Qy      61  CTTCTCTGCTTATTAATGACATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
Db      61  CTTCTCTGCTTATTAATGACATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
Qy     121  CTCATGTTTCACATTCCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
Db     121  CTCATGTTTCACATTCCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
Qy     181  CTGACCGAGGCCCATAGACACCTTTGCGGACACCATCAAGGGCTTTGCGATGCGCA 240
Db     181  CTGACCGAGGCCCATAGACACCTTTGCGGACACCATCAAGGGCTTTGCGATGCGCA 240
Qy     241  CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGCTACCTTAA 300
Db     241  CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGCTACCTTAA 300
Qy     301  TTATTAATCTTTAGATATTAATTAATTTTGAATAAATAACAATTAATCTTTGTGA 360
Db     301  TTATTAATCTTTAGATATTAATTAATTTTGAATAAATAACAATTAATCTTTGTGA 360
Qy     361  GGGCTCAGACATGATTTTCGCTTAAGGGCCAGAAATGCGAGGACCAAGCATCTAGTG 420
Db     361  GGGCTCAGACATGATTTTCGCTTAAGGGCCAGAAATGCGAGGACCAAGCATCTAGTG 420
Qy     421  TTCACATTTGGGCTCCCGACAGATTTAAATAAATAACAAGTAACTTAATCCACT 480
Db     421  TTCACATTTGGGCTCCCGACAGATTTAAATAAATAACAAGTAACTTAATCCACT 480
Qy     481  CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAACACGATCCTTTAAATA 540
Db     481  CGAAGCTATCATGATATGTTTAAAGAAATCTATTTAAACACGATCCTTTAAATA 540
Qy     541  CAAACATTTTGAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAA 600
Db     541  CAAACATTTTGAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAA 600
Qy     601  TTAATATGAAAGGTAGCTATGCGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
Db     601  TTAATATGAAAGGTAGCTATGCGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
Qy     661  ATTGTTTATATATACATTTTCTTCTTAATAGATGATTTTCTTCCGATTTTAA 720
Db     661  ATTGTTTATATATACATTTTCTTCTTAATAGATGATTTTCTTCCGATTTTAA 720
Qy     721  ATGACTATTAAGTCAATTTTATTAAGAGACGCAATGCTGATGATCTGTTCAAAATC 780
Db     721  ATGACTATTAAGTCAATTTTATTAAGAGACGCAATGCTGATGATCTGTTCAAAATC 780
Qy     781  TTTCTGATTTTAAAGAGTGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Db     781  TTTCTGATTTTAAAGAGTGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Qy     841  TTCAAAAAAATTAATGTTTCTTCTTAATAAATAGAAAAATCTTAAGAAAAATAGAT 900
Db     841  TTCAAAAAAATTAATGTTTCTTCTTAATAAATAGAAAAATCTTAAGAAAAATAGAT 900
Qy     901  TGGCAGACTAGCCCTTGAATGTTTCCCAATAAATTAACAATCACTGTGATTAATTTG 960
Db     901  TGGCAGACTAGCCCTTGAATGTTTCCCAATAAATTAACAATCACTGTGATTAATTTG 960
Qy     961  GCGAGCCCATTAATTAATTAACGAACTGAATCGAGCGAAACCAATCGAGCTAT 1020
Db     961  GCGAGCCCATTAATTAATTAACGAACTGAATCGAGCGAAACCAATCGAGCTAT 1020
Qy    1021  TTCTCTAGATTAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db    1021  TTCTCTAGATTAATAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy    1081  TGAATGTGCGGTTTGGCAACGATACCAACCTTAATCATAGCTCATAGGTGCTTAC 1140
Db    1081  TGAATGTGCGGTTTGGCAACGATACCAACCTTAATCATAGCTCATAGGTGCTTAC 1140

```

```

Db      1081  TGAATGTGCGGTTTGGCAACGATACCAACCTTAATCATAGCTCATAGGTGCTTAC 1140
Qy      1141  GGTTCGAGGCTCTGCTGATCTCAATGAGATCTCAATCTTGTAAACGTTCTGTC 1200
Db      1141  GGTTCGAGGCTCTGCTGATCTCAATGAGATCTCAATCTTGTAAACGTTCTGTC 1200
Qy     1201  TTGTTTCATGCTCAAGCCTTGCCTATTTCTGAACCAAGAGATTAATCTCCAAACAAT 1260
Db     1201  TTGTTTCATGCTCAAGCCTTGCCTATTTCTGAACCAAGAGATTAATCTCCAAACAAT 1260
Qy     1261  CCATCTTAATCAATGCACTTCCATGCAACACGCAATATGTTTCTGAACAGATCTATT 1320
Db     1261  CCATCTTAATCAATGCACTTCCATGCAACACGCAATATGTTTCTGAACAGATCTATT 1320
Qy     1321  AAAGATCAACAGCTAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGGATCTTTT 1380
Db     1321  AAAGATCAACAGCTAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGGATCTTTT 1380
Qy     1381  CGTCCACGACCATG 1394
Db     1381  CGTCCACGACCATG 1394

```

```

RESULT 2
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFPMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/680,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

```

```

Query Match      99.7%; Score 1389.2; DB 20; Length 1394;
Best Local Similarity 99.8%; Pred. No. 1.2e-293;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  CCAATGCTCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Db      1  CCAATGCTCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
Qy     61  CTTCTCTGCTTATTAATGACATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
Db     61  CTTCTCTGCTTATTAATGACATGAAATCGGGGTTACAAAAAATCTTCCAGGGTGCATGAT 120
Qy     121  CTCATGTTTCACATTCCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
Db     121  CTCATGTTTCACATTCCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
Qy     181  CTGACCGAGGCCCATAGACACCTTTGCGGACACCATCAAGGGCTTTGCGATGCGCA 240
Db     181  CTGACCGAGGCCCATAGACACCTTTGCGGACACCATCAAGGGCTTTGCGATGCGCA 240
Qy     241  CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGCTACCTTAA 300
Db     241  CGAGAGTATCGGGTGTGTGATCCAGGGGATATATGTCGCCCAATGCTACCTTAA 300
Qy     301  TTATTAATCTTTAGATATTAATTAATTTTGAATAAATAACAATTAATCTTTGTGA 360
Db     301  TTATTAATCTTTAGATATTAATTAATTTTGAATAAATAACAATTAATCTTTGTGA 360

```


/ LENGTH: 158
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-3

Query Match 10.5%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 5.2e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGGTGATCTGACATGAGCACTGCTGTTCAACCGTGTGTC-TTGTTCCATGTC 1213
|||||
DB 1 CGGTGATCTGACATGAGCACTGCTGTTCAACCGTGTGTTGTTCCATGTC 60

QY 1214 CAAGCTTGCTATTTCTGTAACCAAGAGATACCTACTCCCAACAATCTTACTCAT 1273
|||||
DB 61 CAAGCTTGCTATTTCTGTAACCAAGAGATACCTACTCCCAACAATCTTACTCAT 120

QY 1274 GCACTTCCATGCAAAACGACATATGTTTCTGTAAC 1311
|||||
DB 121 GCACTTCCATGCAAAACGACATATGTTTCTGTAAC 158

RESULT 5
US-10-425-115-134230

/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B

/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230

/ LENGTH: 1261
/ TYPE: DNA

/ ORGANISM: Zea mays
/ FEATURE:

/ OTHER INFORMATION: Clone ID: MR14577_53903C.1
US-10-425-115-134230

Query Match 5.8%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 2.7e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 792 TTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
|||||
DB 999 TCTTAAGGCTAGTTTGAAGACCTTATTTTAAAGATTTCTTTTAAAGTAA 1058

QY 852 TTAGTTATTTTCTCTTTTAAATGAAAACCTTGAAGAAATAGAGTTGCCAGACTAG 911
|||||
DB 1059 TTAGTTATTTTCTCTTTGAGAAATGAAATCTTTGAGAAATGAGTTGCTAAACTAG 1118

QY 912 CCT 915
|||||
DB 1119 CCT 1122

RESULT 6
US-10-739-930-1857

/ Sequence 1857, Application US/10739930
/ Publication No. US20040216190A1

/ GENERAL INFORMATION:

/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

/ FILE REFERENCE: 38-21(53377)B
/ CURRENT APPLICATION NUMBER: US/10/739,930

/ CURRENT FILING DATE: 2003-12-18

/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 1857
/ LENGTH: 1663
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match 5.5%; Score 76.2; DB 20; Length 1663;
Best Local Similarity 96.3%; Pred. No. 3.1e-06;
Matches 78; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1314 ATCTATTAAGATCAACAAGCTAGCTTCTCCGCTAGTCTCCCTCTCTGCGCA 1373
|||||
DB 2 ATCTATTAAGATCAACAAGCTAGCTTCTCCGCTAGTCTCCCTCTCTGCGCA 61

QY 1374 TCTTTTGTGTCACCAACCATG 1394
|||||
DB 62 TCTTTTGTGTCACCAACCATG 82

RESULT 7
US-10-425-114-23340

/ Sequence 23340, Application US/10425114
/ Publication No. US2004003488A1

/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E

/ APPLICANT: Tabaska, Jack E
/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53313)B

/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340

/ LENGTH: 320
/ TYPE: DNA

/ ORGANISM: Zea mays
/ FEATURE:

/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340

Query Match 5.3%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.2e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 791 TTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
|||||
DB 62 TTTTAAGGCTAGTTTGAAGAAATCTCATTTTTCAGAGATTTTATTTTCTTAAAGGA 121

QY 851 ATTAGTTATTTTCTCTTATTAATAGAAAACCTTGAAGAAATAGAGTTGCCAGACTA 910
|||||
DB 122 ATTAGTTATTTTCTCTTGAAGAAATAGAAATTCCTTGGAAATAGAGTTCTTAAAGCA 181

QY 911 GCCCTGAAT 920
|||||
DB 182 GCCCTGAAT 191

RESULT 8
US-10-425-114-16264

/ Sequence 16264, Application US/10425114
/ Publication No. US2004003488A1

/ GENERAL INFORMATION:

/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua

/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E


```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313) B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16264
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264
```

```

Query Match
Best Local Similarity 73.1%; Score 74; DB 18; Length 624;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 791 TTTTAAGAGCTAGTTGGCAACCCGTGTTCTTTCAAGAAATTTTGATTTTTCAGAAAAA 850
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 374 TTTTGGGCTAGTTGGAATCTATTTTTCAGAGATTTTATTTTCTTAAGAA 433
QY 851 ATTAGTTATTTCTCTTTATTAATAAGAAACACTAGAAAAATAGAGTGCAGACTA 910
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 434 ATTAGTTATTTTCTTTTGAGAAATAGAAATCCCTTGGAAATAGAGTTCTAAACGA 493
QY 911 GCCCTAGAAAT 920
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 494 GCCCTTAATT 503
```

```

RESULT 9
US-10-425-115-141826
; Sequence 141826, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 141826
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826
```

```

Query Match
Best Local Similarity 69.1%; Score 73; DB 20; Length 1326;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
```

```

QY 756 TGTCTGAGATTCCTCGTTCAAAAATCTTGTGATTTTCTTAAGAGCTAGTTGGCAACCT 815
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1090 TGAAGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1149
QY 816 GTTCTTTCAAGAAATTTGATTTTTCAGAAAAAATGATTTTCTCTTATATAA 875
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1150 ATTT-TTCCACTGATTTTCTTCTTAAGAAATAGTATTTCCCTTGGAGAAA 1208
QY 876 TAGAAAACACTTAGAAAAATAGAGTGCAGACTAGCCCTAGAAAT 920
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1209 TAGGAATCCCTTAGAAAAAATAGTTCCAACTAGCCCTTAAT 1253
```

```

RESULT 10
US-10-425-114-32493
; Sequence 32493, Application US/10425114
; Publication No. US2004003488B1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313) B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32493
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493
```

```

Query Match
Best Local Similarity 73.6%; Score 72.8; DB 18; Length 2445;
Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
```

```

QY 792 TTTAAGAGCTAGTTGGCAACCCGTGTTCTT-TCAAGAAATTTTGATTTTTCAGAAAAA 850
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1623 TCTTGGGCTAGTTGAGAACCCCTTTTTCAGAAATTTTCATTTTTCAGAAAAA 1682
QY 851 ATTAGTTATTTCTCTTTATTAATAAGAAACACTAGAAAAATAGAGTGCAGACTA 910
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1683 ATTAGTTATTTTCTCTTGAAGAAATAGAAATCCCTTAATAATAATAGTGTTCAGAACTA 1742
QY 911 GCCCTAGAAATGTTTCCCAATAA 934
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1743 GCCCTTAATGTTTTCATGAA 1766
```

```

RESULT 11
US-10-425-115-83293/C
; Sequence 83293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 83293
; LENGTH: 2729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2729)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_175978C.1
US-10-425-115-83293
```

```

Query Match
Best Local Similarity 60.7%; Score 72.8; DB 20; Length 2729;
Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
```


Publication No. US20040034889A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 31957
LENGTH: 2537
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73238H12_FLI
US-10-425-114-31957

Query Match 5.2%; Score 71.8; DB 18; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.5e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
QY 792 TTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAAA 851
DB 1799 TCTTAGGGCTAGTTGGAACCT-TTTTCCCAAAAGATTTTCATTTTCMAAGAAAA 1857
QY 852 TTAGTTATTTCTCTTAATAAAGAAACCTAGAAATAGAGTTGCCAGACTAG 911
DB 1858 TTAGTTCATTTTCTTGAAGAAATAGGAATCCCTTAATAAATAGTGTTCMAACTAG 1917
QY 912 CCTAGATGTTTCCCAATAAA 934
DB 1918 CCTTAATGTTTTCATGAA 1940

Search completed: September 15, 2005, 20:45:53
Job time : 908.859 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 : Search time 4347.69 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2

Perfect score: 1394

Sequence: 1 cccttggtgctcctcgaaga.....cttttcgctccaccaccatg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_eat1:*
2: gb_eat2:*
3: gb_hlc:*
4: gb_eat3:*
5: gb_eat4:*
6: gb_eat5:*
7: gb_eat6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	948.8	68.1	963	9	CC656933 OGDQ20TM
2	755.6	54.2	915	9	CG224225 OG1AG08TV
3	496.4	35.6	687	9	CC656939 OGDQ20TV
4	96	6.9	715	9	CG252571 CGAB805TC
5	92	6.6	967	9	CL235046 ZMMBB057
6	88	6.3	814	9	CG048704 PU1019TB
7	87.8	6.3	754	9	CG414922 ZMMBB029
8	86.8	6.2	950	8	CC439901 PUHRI5TB
9	86.6	6.2	652	8	CC384247 PUHOC7TB
10	86.6	6.2	797	8	CC400575 PUHUI61TD
11	86.6	6.2	820	8	CC400574 PUHUI61TB
12	85.4	6.1	471	9	CG103452 PUHBI9TB
13	85.2	6.1	765	9	CG082135 PUHOK12TD
14	85.2	6.1	781	9	CG349565 OG0F183TH
15	85.2	6.1	815	9	CG349565 OG0F183TH
16	85.2	6.1	834	9	CG349565 OG0F183TH
17	85	6.1	789	8	CC433618 PUHHP17TD
18	84.6	6.1	1092	8	CNS020K7
19	84	6.0	793	8	BZ816381
20	84	6.0	1078	9	CL997678 ZMMBH701
21	83.8	6.0	530	9	CG201774 PUIC24TB
22	83.6	6.0	722	9	CG333914 OG0AD14TH
23	83.6	6.0	722	9	CG333929 OG0AD14TV
24	83.6	6.0	781	9	CG034985 PUHGR68TB

25	83.6	6.0	861	8	BZ797976
26	83.6	6.0	865	8	CC430754
27	81.8	5.9	861	9	CG102052
28	81.8	5.9	947	8	CC435780
29	81.8	5.9	981	8	BZ784278
30	81.8	5.9	1017	9	CC620594
31	81.6	5.9	649	9	CC613918
32	81.6	5.9	733	8	BZ778636
33	81.6	5.9	925	9	CG071791
34	81.6	5.9	960	8	BZ676889
35	81.6	5.9	994	8	CC003943
36	81.6	5.9	1016	9	CL996481
37	81.6	5.9	1022	9	CL984151
38	81.4	5.8	999	8	CC385762
39	81.2	5.8	756	9	CG221653
40	81	5.8	1101	9	CNS001R2
41	80.6	5.8	699	8	BZ96930
42	80.6	5.8	832	9	CC676387
43	80.6	5.8	839	9	CG246849
44	80.6	5.8	843	9	CG254365
45	80.6	5.8	852	8	CC385881

ALIGNMENTS

RESULT 1
LOCUS OGDQ20TM 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGDQ20TM ZM 0.7-1.5 KB Zea mays genomic clone ZMMMA0554D15,
CC656933 genomic survey sequence.

ACCESSION CC656933

VERSION CC656933.1 GI:32060225

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robb,D. and Lakey,N.
1 (bases 1 to 963)

TITLE
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGDQ20TV
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMMA0554D15"
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 68.1% Score 948.8; DB 9; Length 963;
Best Local Similarity 99.7%; Pred. No. 1.5e-190;
Matches 961; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
404 ACCAGCATGTCTAGTGTCTACTATTTGACATACCAAGAAAGATTAAATTAAC 463
1 ACCAGCATGTCTAGTGTCTACTATTTGACATACCAAGAAAGATTAAATTAAC 60


```

QY 464 AAGTAATCACTACGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACC 523
    |||
    |||
    |||
Db 61 AAGTAATCACTACGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACC 120
QY 524 AGCATCTCTTAAACAAAGCATATTTGAAAAGACAAATATGTTACAGTTTACAA 563
    |||
    |||
    |||
Db 121 AGCATCTCTTAAACAAAGCATATTTGAAAAGACAAATATGTTACAGTTTACAA 180
QY 584 CATCTAAGACGCAAAATTAATCGAAAGTAGCTAGCGTTGAGATTTTCTTTTC 643
    |||
    |||
    |||
Db 181 CATCTAAGACGCAAAATTAATCGAAAGTAGCTAGCGTTGAGATTTTCTTTTC 240
QY 644 ATTCTGTATTTTGTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 703
    |||
    |||
    |||
Db 241 ATTCTGTATTTTGTATGTTTATATACATTTTCTCTTCAATAGAGTATTT 300
QY 704 TCTTCGATTTTAAATGACTATAAAGTATTTTATATAAAGACGAGTGTGTA 763
    |||
    |||
    |||
Db 301 TCTTCGATTTTAAATGACTATAAAGTATTTTATATAAAGACGAGTGTGTA 360
QY 764 ATTCTGTCAAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTT 823
    |||
    |||
    |||
Db 361 ATTCTGTCAAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTT 420
QY 824 CAAAGATTTGATTTTTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAC 883
    |||
    |||
    |||
Db 421 CAAAGATTTGATTTTTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAC 480
QY 884 ACTTGAAGAAATAGAGTGGCAGACTAGCCCTGAAATGTTTCCCAATAATTAATCA 943
    |||
    |||
    |||
Db 481 ACTTGAAGAAATAGAGTGGCAGACTAGCCCTGAAATGTTTCCCAATAATTAATCA 540
QY 944 CTGTGTATATTTTGGCAGGCCATTAATTTAAACGAACTGAAATCGAGCA 1003
    |||
    |||
    |||
Db 541 CTGTGTATATTTTGGCAGGCCATTAATTTAAACGAACTGAAATCGAGCA 600
QY 1004 AACCAATCTGACCTATTTCTTGAATTAATAAAGGAGAGAGAGAGAGAGAGAGAG 1063
    |||
    |||
    |||
Db 601 AACCAATCTGACCTATTTCTTGAATTAATAAAGGAGAGAGAGAGAGAGAGAG 660
QY 1064 TTTTAAAGTCTGTCTGAGATGTGGGTTTGGCAACGATAGCCACCTGATCATAGCT 1123
    |||
    |||
    |||
Db 661 TTTTAAAGTCTGTCTGAGATGTGGGTTTGGCAACGATAGCCACCTGATCATAGCT 720
QY 1124 CATAGGCTAGTGAAGTGGGAGCTCTGATCATCTGATAGGCACTATCATATGC 1183
    |||
    |||
    |||
Db 721 CATAGGCTAGTGAAGTGGGAGCTCTGATCATCTGATAGGCACTATCATATGC 780
QY 1184 TTGTTCAACCGTCTGTTGTTCCATGTTCAAGCCCTGCTATTTGAAACCAAGAGAT 1243
    |||
    |||
    |||
Db 781 TTGTTCAACCGTCTGTTGTTCCATGTTCAAGCCCTGCTATTTGAAACCAAGAGAT 840
QY 1244 ACTTACTCCCAAAACATCTTACTCATGCAATTCATGCAAAACGACATATGTT 1303
    |||
    |||
    |||
Db 841 ACTTACTCCCAAAACATCTTACTCATGCAATTCATGCAAAACGACATATGTT 900
QY 1304 TCTGTAACGATATTTAAAGATCAACAAGCTAGAGGTTCTCCGCTAGTTCCCTCTC 1363
    |||
    |||
    |||
Db 901 TCTGTAACGATATTTAAAGATCAACAAGCTAGAGGTTCTCCGCTAGTTCCCTCTC 959
QY 1364 CCTC 1367
    |||
    |||
    |||
Db 960 CCTC 963

```

```

RESULT 2
CG224225
LOCUS CG224225 915 bp DNA linear GSS 22-AUG-2003
DEFINITION OGIAG08TV ZM_0.7-1.5_KB Zea mays genomic clone ZMMBMA0716B15,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113

```

```

KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 915)
AUTHORS Whiteley, C.A., Quackenbush, J., Van Aken, S., Utecher, T.,
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
CONSORTIUM Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGIAG08TV
Contact: Cathy Whiteley
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteley@cigr.org
Seq primer: TP
Class: sheared ends.
FEATURES
source location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.7-1.5_KB"
/clone="ZMMBMA0716B15"
/notes="vector: pBCSK-1; Site: 1; HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 54.2%; Score 755.6; DB 9; Length 915;
Best Local Similarity 99.4%; Pred. No. 1.3e-149;
Matches 769; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY 621 TGACGTTGAGATTTTCTTTTTCATTTCTGTTATTTGTTTATTAATTAATTTT 680
    |||
    |||
    |||
Db 1 TGACGTTGAGATTTTCTTTTTC-TTCTGTTATTTGTTTATGTTTATTAATTAATTTT 59
QY 681 CTTCCTTCAATTAAGATGATTTTCTTCGATTTTATAAATGACTATTAAGTCAATTTT 740
    |||
    |||
    |||
Db 60 CTTCCTTCAATTAAGATGATTTTCTTCGATTTTATAAATGACTATTAAGTCAATTTT 119
QY 741 AATTAAGGACGACATGTCGATATTTCTGCAAAATCTTTCGATTTTATTAAGGC 800
    |||
    |||
    |||
Db 120 AATTAAGGACGACATGTCGATATTTCTGCAAAATCTTTCGATTTTATTAAGGC 179
QY 801 TAGTTGGCAACCGTCTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTAT 860
    |||
    |||
    |||
Db 180 TAGTTGGCAACCGTCTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTAT 239
QY 861 TTTCTCTTATTAATAATGAAAACTTAAGAAAAATAGAGTGGCAGACTAGCCCTAGAT 920
    |||
    |||
    |||
Db 240 TTTCTCTTATTAATAATGAAAACTTAAGAAAAATAGAGTGGCAGACTAGCCCTAGAT 299
QY 921 GTTTTCCCAATTAATTAACATCTGATTAATTAATTTGGCAGCCCATTAATTTAT 980
    |||
    |||
    |||
Db 300 GTTTTCCCAATTAATTAACATCTGATTAATTAATTTGGCAGCCCATTAATTTAT 359
QY 981 AAACGAAACTGAATTCAGAGCAAAACCAATCTGAGTATTTCTGATTAATTAAG 1040
    |||
    |||
    |||
Db 360 AAACGAAACTGAATTCAGAGCAAAACCAATCTGAGTATTTCTGATTAATTAAG 419
QY 1041 GGAGAGAGAGAGAGAAATCAATTTTAAGTATTTGCTCGAGATGTGCGGTTGGCAA 1100
    |||
    |||
    |||
Db 420 GGAGAGAGAGAGAGAAATCAATTTTAAGTATTTGCTCGAGATGTGCGGTTGGCAA 479
QY 1101 CGATAGCAACCGTATCATAGTCAATAGTGTCTGATGCTGAGTGGCAGCTGCTGTC 1160
    |||
    |||
    |||
Db 480 CGATAGCAACCGTATCATAGTCAATAGTGTCTGATGCTGAGTGGCAGCTGCTGTC 539

```



```
QY 1161 ATCTACATGSCATACACTAGCTGTTGTTCAACCGTTGCTGTTGCTGCTCAAGCCT 1220
    |||||
Db 540 ATCTCATATGGCATACTACATGCTGTTGTTCAACCGTTGCTGTTGCTGCTCAAGCCT 599
QY 1221 TGGCTATTCTGAACCAAGAGATACCTACTGCCAACAATCATCTTACTCATGCAACT 1280
    |||||
Db 600 TGCGTATTCTGAACCAAGAGATACCTACTGCCAACAATCATCTTACTCATGCAACT 659
QY 1281 CCATGCAAAACGACCATATGTTTCTGTAACAGATCTATTAAAGATACAAACGCTAGCG 1340
    |||||
Db 660 CCATGCAAAACGACCATATGTTTCTGTAACAGATCTATTAAAGATACAAACGCTAGCG 719
QY 1341 TTCTCCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1394
    |||||
Db 720 TTCTCCGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 773

RESULT 3
CC656939/c 687 bp DNA linear GSS 19-JUN-2003
LOCUS OGMDQ20TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM54D15,
DEFINITION genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 687)
White, C.A., Quackenbush, J., Van Aken, S., Uteback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGMDQ20TM
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
    1..687
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZM54D15"
    /clone_1lb="ZM_0.7_1.5_KB"
    /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
Query Match 35.6%; Score 496.4; DB 9; Length 687;
Best Local Similarity 98.8%; Pred. No. 1.1e-94;
Matches 500; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 889 GAAAAATAGAGTTGCCAGACTAGCGCTGAGATGTTTCCCAATTAATTACATCACTGTG 948
    |||||
Db 687 GAAAAATAGAGTCCCGAGCTAGCGCTGAGATGTTTCCCAATTAATTACATCACTGTG 628
QY 949 TATAATATTATTTGGCGAGCCCATTAATAATTAAACGAAATCGAATCGAGGAACCA 1008
    |||||
Db 627 TATAATATTATTTGGCGAGCCCATTAATAATTAAACGAAATCGAATCGAGGAACCA 568
QY 1009 AATCTGAGTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTGA 1068
    |||||
Db 567 AATCTGAGTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTGA 508
```

```
QY 1069 AGTCATTGTCCTGAGATGAGCGGTTTGGCAACGATAGGCACCGTATCATAGTCATAG 1128
    |||||
Db 507 AGTCATTGTCCTGAGATGAGCGGTTTGGCAACGATAGGCACCGTATCATAGTCATAG 448
QY 1129 GTGCTACGTCAGGTTGCGAGCTCTGTCATCTCATATGATGATGATGATGATGATGAT 1188
    |||||
Db 447 GTGCTACGTCAGGTTGCGAGCTCTGTCATCTCATATGATGATGATGATGATGATGAT 388
QY 1189 CAACGCTGCTGTTGTTCCATGTCACAGCTTCCCTATTCTGAAACCAAGAGATACCTA 1248
    |||||
Db 387 CAACGCTGCTGTTGTTCCATGTCACAGCTTCCCTATTCTGAAACCAAGAGATACCTA 328
QY 1249 CTCGCAAAACATCATCTACTCATGCAACTTCATGCAAAACGACATATGTTTCCG 1308
    |||||
Db 327 CTCGCAAAACATCATCTACTCATGCAACTTCATGCAAAACGACATATGTTTCCG 268
QY 1309 AACGATCTATTAAAGATACAAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCT 1368
    |||||
Db 267 AACGATCTATTAAAGATACAAACAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCT 208
QY 1369 GCCGATCTTTTTCGTCACACCATG 1394
    |||||
Db 207 GCCGATCTTTTTCGTCACACCATG 182

RESULT 4
CG252571/c 715 bp DNA linear GSS 25-AUG-2003
LOCUS OG4B805TC_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZM54D10,
DEFINITION genomic survey sequence.
ACCESSION CG252571
VERSION CG252571.1 GI:34154661
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 715)
White, C.A., Quackenbush, J., Van Aken, S., Uteback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
    1..715
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /strain="B73"
    /db_xref="taxon:4577"
    /clone="ZM54D10"
    /clone_1lb="ZM_0.7_1.5_KB"
    /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
    methylation filtered genomic DNA library"

ORIGIN
Query Match 6.9%; Score 96; DB 9; Length 715;
Best Local Similarity 80.1%; Pred. No. 8.3e-10;
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 135 TCTCCACCTCGCGTTGACATTTCTTGATGTCGGTGGTCCATCTGACCGAGGCCCA 194
    |||||
Db 356 TCACACACCGGCTTGTGCTTCTTGATGTCGGTGGTCCATCTGACCGAGGCCCA 297
QY 195 TCACACACCTTTCGGACACCATCAAGGCGCTTTCGATGCGCCACGACGATCGGG 254
```


Db 296 TAAGGACCTGTCGGAGAACCACTTAAGGCTTTGGATGGCCAGGACAT-CGGGG 238
QY 255 TCCTGTGATCCAGGAGATATATGTCCCCCAACATC 290
Db 237 TCGTGTGACCCGGGGATATATATCCCAACACC 202

RESULT 5
CL235046 967 bp DNA linear GSS 15-JAN-2004
LOCUS ZMBBB0575001.r ZMBBB (HindIII) Zea mays genomic clone
DEFINITION ZMBBB0575001 3', genomic survey sequence.
ACCESSION CL235046
VERSION CL235046
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 967)
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
Sequencing of the maize genome at PGR (2003c)
Unpublished (2003)
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 64.
Location/Qualifiers
1..967
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMBBB0575001"
/lab_host="E. coli DH10B"
/clone_lib="ZMBBB (HindIII)"
/note="vector: pCUG1; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 6.6%; Score 92; DB 9; Length 967;
Best Local Similarity 78.6%; Pred. No. 5.9e-09;
Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 150 TGCACATTTCTTGAGATGCGTGTGTCATCTGACCGAGGCCCATCAGACCTTTCCG 209
Db 827 TGGCGCTTCTTGAGATCCGGCGGTTCCGCTGATGAGGCCCATTCAGACCTTTTGG 886
QY 210 GACACCATCAAGGCGCTTTCGATGCGCCACGAGACGTATCGGTCGTGTATCCAGG 269
Db 887 GTGGCCCATCAAGTGCCTTTCGATGTCCACGGAGATGATCCGGGCCGTGTGACCTGGG 946
QY 270 GGATATATGTCCCCCAACAT 289
Db 947 GGATATCTGTCCCCCATAGT 966

RESULT 6
CG048704 814 bp DNA linear GSS 19-AUG-2003
LOCUS PUILO19TB_ZM_0.6_1.0_KB Zea mays genomic clone ZMBBB0611C13,
DEFINITION genomic survey sequence.
ACCESSION CG048704
VERSION CG048704.1 GI:33920884

KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 814)
WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUILO19TB
Contact: Cathy WhiteIaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..814
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBBB0611C13"
/clone_lib="ZM_0.6_1.0_KB"
/note="vector: pCR4-toPO; Site_1: EcoRI; 0.6-1.0 kb high
cor selected genomic DNA library"

ORIGIN

Query Match 6.3%; Score 88; DB 9; Length 814;
Best Local Similarity 62.7%; Pred. No. 4.1e-08;
Matches 153; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTTAAAGCATTTTATATAGAGCAGCATGTCGTATTCGT 772
Db 487 TTTAGAGACGTTGTGATGATGAGGAAGATATAGAAATATATCTTTAGGAATGTA 546
QY 773 CAAAAATCTTCTGATTTTATTAAGAGCTAGTTGGCAACCTGTCTTTCAAGAAAT 832
Db 547 TAAAGGACAGAAACATTTCTTAAGAGCTAGTTGTATCCCATTTTTCAAAGAT 606
QY 833 TTGATTTTTCAAAA-AAATTAAGTTTATTTCTCTTATTAATAAAGAAAACATTAGAA 891
Db 607 TTCATTCTCACAAAGGAAAAATATGTTTATTTCCCTTTGTAAATGTATATCATCTCGAA 666
QY 892 AATAGAGTTGCGACAGTACCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTAT 951
Db 667 AATATAGATTCCTCAAACTAGCCCTTAAGGTATATCCCATTAATCCCTATTCATTTCT 726
QY 952 AATT 955
Db 727 ATTT 730

RESULT 7
CG414922 754 bp DNA linear GSS 08-SEP-2003
LOCUS ZMBBB0290L09.r ZMBBB Zea mays genomic clone ZMBBB0290L09 3',
DEFINITION genomic survey sequence.
ACCESSION CG414922
VERSION CG414922.1 GI:34505144
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 754)
Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.

	Matches	131,	Conservative	0;	Mismatches	74;	Indels	0;	Gaps	0;
Qy	714	TTATATAATGAC	TATAAAGTCATTTT	TATATATAGACACG	ATGTCGTAGATTTCTG	TT	773			
Db	341	TAAATTAATTTG	AGTACGTACGTACATCTT	TAATTCGACATATTTATG	TATGTTAAATTAATGCTA	400				
Qy	774	AAAAATCTTCG	ATTTTTTTTAAAGAGTCAG	TTGGCAACCCGTTCTTT	CAAGATTT	833				
Db	401	AGAACTTATTT	TATTAATATGTTAAGAGCTAG	ATTTGGCAACTATATTTT	CTAAGACATTTT	460				
Qy	834	TGATTTTTTCA	AAAAAAATTAAGTTATTTT	CTTTATATAATGAA	AACACTTAGAAA	893				
Db	461	TATTTTTTCA	AGTAAATTAAGTTATTTT	TCATGGGAAAAATAGAA	TTATTTGAAAA	520				
Qy	894	ATAGAGTTCC	AGACTAGCCCTAGA	918						
Db	521	ATTAGTTTCA	AACTAGCCCTAAA	545						

RESULT	10
CC400575/c	
LOCUS	CC400575
DEFINITION	797 bp DNA linear GSS 19-MAY-2003
ACCESSION	PJRLB61TD ZM 0.6.1_0_KB Zea mays genomic clone ZMMBRta480L01,
VERSION	CC400575
KEYWORDS	genomic survey sequence.
SOURCE	CC400575.1 GI:30880665
ORGANISM	GSS.
	Zea mays
	Zea mays

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 797)	Whitlaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Benneken, J.	Maize Genomics Consortium	Unpublished (2003)	
Other GSSs:	PGHLU61TB			

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PURL51TB
Contact: Cathy White|aw
TIGR 7112 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white|aw@tigr.org
Seq Primer: TP
Class: sheared ends.

FEATURES	Location/Qualifiers
source	1. .797

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBT480L01"
/clone_1lb="2M 0.6-1.0 KB"
/note="Vector: pCR4::TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

```

ORIGIN

Query Match	6.2%	Score 86.6	DB 8	Length 797
Best Local Similarity	59.7%	Pred. No. 8.2e-08		
Matches 163; Conservative	0	Mismatches 109	Indels 1	Gaps 1

713 TTTATAAATGACTATTAAGTCATTTTATATAGAAGCAGCATGTCTGATTTCTCGTT 772

Db 577 TTITAGAGACGTTGTTAGATGGAAGAATATAGAGAAAATATCTTTTITAGAAATGTTA 518

773 CAAAAATCTTCTGATTTTTTTTAAAGAGCTAGTTGGCAACCGTTCTTTTCAAGAATT 832

Db 517 TAAAGCAAGAAATTTTAAAGCTAGTTGGTAATCCATTTTTCAAAGATT 458

833 TTGAATTTTCAAAA-AAAATTAGTTATTTCTCTTAATAAAATGAAAAACACTTAGAA 891

Db 457 TTCATTTCCACAAGGAAAAAATGATTTTATTTTCCCTTTGGTAAATAAGTATACCTGGAA 398

Qy 892 AAATGAGTTGGCAGACATGAGCCCTGAGAAATGTTTTCCCAATAATATACATCATCTGTAT 951

Db 397 AAATGAGTTCTCAAACTGAGCCCTTAAGGTATATCATGATAATATCCCTATTCATTTCTCT 338

Qy 952 AATTATTTGGCCAGCCCATTAATATTTTAAAC 984

Db 337 ATTTCAAATTTCCACTTTGCAAAATGTGTCAATC 305

RESULT 11	
CC400574	
LOCUS	820 bp DNA
DEFINITION	linear GSS 19-MAY-2003
	PubMed17B ZM 0.6 1.0 KB Zea mays genomic clone ZMMB17A480L01,
	genomic survey sequence.

ACCESSION	CC400574	GI:30880664
VERSION	CC400574.1	
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	

REFERENCE
Whitebaw, C.A., Quackenbush, J., Van Aken, S., Uettermack, T.,
1 (pages 1 to 820)

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other GSSs: PHLU61ND
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

9112 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white.law@icigr.org
Seq primer: TR
Class: sheared ends.

```

FEATURES
    SOURCE
        location/Qualifiers
            1. 820
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMBA4601.01"
                /clone_lib="ZM_0.6_1.0_KB"
                /note="vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                COT selected genomic DNA library"

```

ORIGIN

Query Match	6.2%	Score 86.6;	DB 8;	Length 820;
Best Local Similarly	59.7%	Pred. No. 8.2e-08;		
Matches 163; Conservative	0;	Mismatches 109;	Indels 1;	Gaps 1;

713 TTATATAATGACTATAAGTCATTTTATATATAGAGCAGCATGTCGTAGATTCTGTT 772

Db 374 TTTAGGCGCTTGGTAGAGATGGAAGATATAGAGAAATAAATCTTTAGAGATGTTA 433

773 CAAAATCTTTCGATTTTAAAGAGCTAGTTGGCAACCCGTTCCTTCAAGAATT 832

Db 434 TAAAGCAAGAACAATTTTAAAGAGCTAGTTGGTAATCCATTTTTCAAAGGATT 493

Qy 833 TTGATTTTTCACAAA-AAAATTGTTTATTTCTCTTTATATAAGAAAACACTTGAA 891

Db 494 TTCAATTCTACCAAGGGAAAAATAGTTTATTTCCCTTTTGTAATAATGTAATCACTCGGA 553

892 AATAGAGTTGCCAGCTAGCCTAGATGTTTCCCATAAATTACATCACGTGTAT 951

DB 554 AAAAAGAGTCTCAAAACGACCCTAAGGTATATCCGAATAAACCCCCCACTCATTCTCTC 613

OE 2 AATTCTTCCTCCCTGCGGCCCATTAATAATTCTTTGGTAATC 984

[illegible]

Db 614 ATTCAATTCACCTTTCAGAAATAGTGCATC 646

RESULT 12 471 bp DNA linear GSS 20-AUG-2003
 CG103452
 LOCUS PUJBE19TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBR0628D13,
 DEFINITION genomic survey sequence.
 ACCESSION CG103452
 VERSION CG103452.1 GI:33985746
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 471)
 White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennettzen, J.
 Maize Genomics Consortium
 Unpublished (2003)
 Other GSSs: PUJBE19TD
 Contact: Cathy White,law
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: white,law@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source Location/Qualifiers
 1..471
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMBR0628D13"
 /clone_1lb="ZM_0.6_1.0_KB"
 /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"

ORIGIN
 Query Match 6.1%; Score 85.4; DB 9; Length 471;
 Best Local Similarity 69.5%; Pred. No. 1.4e-07;
 Matches 116; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 789 TTTTAAAGAGTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTCAAAA 848
 |||||
 Db 216 TTTTCTTAGAGCTAGTTGGAAATCCCATTTTTCAGGATTTACATTTTTCAGGG 275
 |||||
 QY 849 AATAGTATTATTTCTCTTATTAATAAGAAAACTTAGAAAAATAGTTGCCAGAC 908
 |||||
 Db 276 AATATGATTCATTTCTTCTTGAATAATGGAATCATCGGAAATAGTTCCCAAC 335
 |||||
 QY 909 TAGCCCTAGAAATTTTCCCAATTAATTAATCACTGTGTAAATT 955
 |||||
 Db 336 TAGCGCTTAGACTATTTTCTTAAAAAAGTAGATCAATGATATAGATT 382
 |||||

RESULT 13
 CG082135/c 765 bp DNA linear GSS 20-AUG-2003
 LOCUS PUFQX12TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBR0712B24,
 DEFINITION genomic survey sequence.
 ACCESSION CG082135
 VERSION CG082135.1 GI:33964429
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 765)
 AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennettzen, J.
 TITLE Maize Genomics Consortium
 JOURNAL Unpublished (2003)
 COMMENT Other GSSs: PUFQX12TB
 Contact: Cathy White,law
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: white,law@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source Location/Qualifiers
 1..765
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMBR0712B24"
 /clone_1lb="ZM_0.6_1.0_KB"
 /note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"

ORIGIN
 Query Match 6.1%; Score 85.2; DB 9; Length 765;
 Best Local Similarity 71.8%; Pred. No. 1.6e-07;
 Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 787 ATTTTCTTAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTCAAA 846
 |||||
 Db 255 ATCTTCTAATGACATGATTGGCAACCTTATTT-TCACAGGATTTTATTTTTCAA 197
 |||||
 QY 847 AAAAATTAATTTCTTCTTATTAATAAGAAAACTTAGAAAAATAGTTGCCAG 906
 |||||
 Db 196 GAAATTAATTTTCTTCTTGAATAATTAATCACTTGAGAAAAATATGTTCCAA 137
 |||||
 QY 907 ACTAGCCCTAGATGTTTCCCAATTAATTAATCACTGTGTAAATTATTTTG 960
 |||||
 Db 136 ACTAGCCCTAAGTATTTTATTAATAATAGGAAAAATGTCATTTTTCATTG 83
 |||||

RESULT 14
 CG630219 781 bp DNA linear GSS 19-JUN-2003
 LOCUS OGUG53TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMMA0406J09,
 DEFINITION genomic survey sequence.
 ACCESSION CG630219
 VERSION CG630219.1 GI:32003354
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 781)
 White, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Buddiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGUG53TV
 Contact: Cathy White,law
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: white,law@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 Location/Qualifiers

Query Match	6.1%	Score 85.2	DB 9	Length 781
Best Local Similarity	71.8%	Pred. No. 1.6e-07		
Matches 125, Conservative	0	Mismatches 48	Indels 1	Gaps 1

907 ACTAGCCCTAGATGTTTCCCAATAATTACAATCACTGTGTATAATTATTC 960
546 ACTAGCCCTAAGTATTTTATAAAATAGAGAAATTCGTATTTTCATTG 4933

Accession	Size	Type	Library
CG349565	815 bp	DNA	linear GSS 26-AUG-2003
OG0FI83TH	ZM_0.7_1.5_KB	genomic clone	ZMMBMA0705N21,
		genomic survey sequence.	

Zea mays
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Unpublished (2002)
Other_GSSs: OG0FI83TV

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: white1aw@icgr.org
Seq primer: TR
Class: sheared ends.

```

Location/Qualifiers
1..815
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBA0705N21"
/clone_1lb="ZM 0.7-1.5 Kb"
/notes="Vector: pBCKS-; Site_1: HincII, 0.7-1.5 kb"
methylation filtered genomic DNA library"

```

Match	6.14;	Score 85.2;	DB 9;	Length 815;
Local Similarity	71.84;	Pred. No. 1.6e-07;		
125; Conservative	0;	Mismatches 48;	Indels 1;	Gaps 1

787 ATTTTAAAGAGCTAGTTTGGCAACCTGTTCTTCAAGAATTGATTTTCAAA 846

173 A C T T T C T A A T G A C T A G T T T G G A A C C T A T T T - T C T A C G G A A T T T T A T T T T T C A A T 115
847 A A A A A T T A G T T A T T T T T C T T T A T A A A A T A G A A A C A C T T A G A A A A T A G A G T T G C A G 906
114 G A A A T T A G T T A T T T T T C T T G A G A A A A T A T A A A C A C T T G A A A A T A T A G T T T C C A A 55
907 A C T A G C C C A G A A T G T T T C C A A T A A T T A C A A T C A C T G T G A T A T T A T T T T G 960
54 A C T A G C C C A A A G A T A T T T A T A A A A T A G A G A A A A T G C T A T A T T T T C A T T G 1

Completed: September 16, 2005, 08:08:41
4348.69 secs

us-10-713-381-2_copy_1155_1311.rge

GenCore version 5.1.6
(c) 1993 - 2005 Compu

31 ; Search time 526.823 Seconds
(without statements)

14440.280 Million cell updates/sec

1311

.....cgcacatattgttcctgaac 157

residues

ameters: 9416466

3

predicted by chance to have a score of the result being printed total score distribution.

IES

Description
AX224402 Sequence
AX224394 Sequence
AX224395 Sequence
BD062176 Male tis
BD062177 Male tis
AF360395 Zee mays
AX224396 Sequence
AC135200
Oryza ba
AX224398 Sequence
AX224399 Sequence
EX950854 Danilo re
AC104134 Homo sap
AC136896 Homo sap
CR847532 Danilo re
BX640577 Zebrafish
AC011290 Homo sap
AC091864 Homo sap
AC146134 Pan trog
AC103745 Homo sap

C	20	32.8	20.9	9361.0	9	AC006602	AC006602 Homo sapi
C	21	32.8	20.9	130031	2	AC135264	AC135264 Homo sapi
C	22	32.8	20.9	148598	8	HSBA51C14	AL121875 Human DNA
C	23	32.8	20.9	158316	9	AC135350	AC135350 Homo sapi
C	24	32.8	20.9	169771	9	AC087485	AC087485 Homo sapi
C	25	32.8	20.9	169867	9	AC019229	AC019229 Homo sapi
C	26	32.8	20.9	193559	2	AC149044	AC149044 Pan trogl
C	27	32.6	20.8	53688	2	AC099973	AC099973 Mus muscu
C	28	32.6	20.8	146080	5	AL953855	AL953855 Zebrafish
C	29	32.6	20.8	165845	2	AC137012	AC137012 Rattus no
C	30	32.6	20.8	186829	2	AC147308	AC147308 Pan trogl
C	31	32.6	20.8	204150	2	AC140696	AC140696 Rattus no
C	32	32.6	20.8	252534	2	AC097405	AC097405 Rattus no
C	33	32.4	20.6	2070	10	BC020027	BC020027 Mus muscu
C	34	32.4	20.6	2039	10	BC038066	BC038066 Mus muscu
C	35	32.4	20.6	2443	10	BC055469	BC055469 Mus muscu
C	36	32.4	20.6	110000	1	AE017197_05	Continuation (6 of
C	37	32.4	20.6	137832	2	AC104550	AC104550 Mus muscu
C	38	32.4	20.6	168889	10	AL731779	AL731779 Mouse DNA
C	39	32.4	20.6	182847	2	AC148229	AC148229 Gallus ga
C	40	32.4	20.6	221338	5	AC147437	AC147437 Gallus ga
C	41	32.4	20.6	262640	2	AC139057	AC139057 Mus muscu
C	42	32.4	20.5	139173	9	AC009140	AC009140 Homo sapi
C	43	32.2	20.5	150116	2	AC034132	AC034132 Homo sapi
C	44	32.2	20.5	189323	2	AC150620	AC150620 Callithr
C	45	32.2	20.5	205954	9	AC008546	AC008546 Homo sapi

ALIGNMENTS

LOCUS	AX224402	255 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 9 from Patent WO016097.				
ACCESSION	AX224402				
VERSION	AX224402.1	GI:15554644			
KEYWORDS	.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	1				
TITLE	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.				
JOURNAL	Male tissue-preferred regulatory region and method of using same				
	Patent: WO 0160997-A 9 23-AUG-2001;				
	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
FEATURES	Location/Qualifiers				
source	1..255				
	/organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match	100.0%;	Score 157;	DB 6;	Length 255;	
Best Local Similarity	100.0%;	Pred. No. 1.8e-40;			
Matches 157;	Conservative	0;	Mismatches	0;	Indels
					Gaps
					0;
QY	1	CGNGTCATCTCATATGGCATCTACTACATAGCTTGTTGCAACGTCGCTTGTTCCATCGTC	60		
Db	15	CGGTGATCTCATATGGCATCTACTACATAGCTTGTTGCAACGTCGCTTGTTCCATCGTC	74		
QY	61	AAGCTTGCGCTATTTCTGAACCAAGAGATACCTACTCTCCAAACAATCATCTTATCTCATG	120		
Db	75	AAGCTTGCGCTATTTCTGAACCAAGAGATACCTACTCTCCAAACAATCATCTTATCTCATG	134		
QY	121	CAACTTCATGCAACACGACATATGTTTCCCTGAAC	157		
Db	135	CAACTTCATGCAACACGACATATGTTTCCCTGAAC	171		
RESULT 2					

AX224394
LOCUS AX224394 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160997.
ACCESSION AX224394
VERSION AX224394.1 GI:15554636
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 1 23-AUG-2001.
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 60
DB 1155 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 1214

QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 3
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001.
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 60
DB 1155 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 1214

QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 120

DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 4
BD062176
LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001.
PIONEER HI BRED INTERNATIONAL, INC

COMMENT
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN, PI TIMMY L KENDALL
PC C12N15/62, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC, C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: linear;
FH key Location/Qualifiers

FEATURES
source Location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 60
DB 1155 CGTGCATTCACATGCGATACACTGCTGTTCAACCGTTGCTTGCATCGTCC 1214

QY 61 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAAACCAAGAGATACCTACCTCCCAACAATCATCTTACTCATG 1274

QY 121 CAATTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAATTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 5
BD062177
LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 1394)

AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
 PIONEER HI-BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
 PF 30-OCT-2001
 PR 19-JUN-1998 JP 199504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN,
 PI TIMMY L KENDALL
 PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
 , C07K14/34, C12Q1/68,
 PC A01H5/00
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers.

FEATURES
 source 1..1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 100.0%; Score 157; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 1.9e-40;
 Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTTGTCATCGTCC 60
 |||||
 DB 1155 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTTGTCATCGTCC 1214
 |||||

QY 61 AAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATG 120
 |||||
 DB 1215 AAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATG 1274
 |||||

QY 121 CAACCTTCATGCAAAACGACATATGTTCTCGAAG 157
 |||||
 DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCGAAG 1311
 |||||

RESULT 6
 AF360356 3343 bp DNA linear PLN 12-MAY-2001
 LOCUS Zea mays male fertility protein (Me45) gene, complete cds.
 ACCESSION AF360356
 VERSION AF360356.1 GI:14028756
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 3343)
 Fox T.W., Trimnell,M.R. and Albertsen,M.C.
REFERENCE Cloning of Me45, a gene required for male fertility from Zea mays
AUTHORS Unpublished
TITLE 2 (bases 1 to 3343)
JOURNAL Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
REFERENCE Direct Submission
AUTHORS Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
JOURNAL H1-Bred Intl. Inc., 7500 N.W. 62nd Ave., P.O. Box 1004, Johnston,
 IA 50131-1004, USA
 Location/Qualifiers
 source 1..3343
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /culivar="B73"
 /db_xref="taxon:4577"
 /chromosome="9L"
 <1392..>2942
 /gene="Me45"
 join(<1392..1768,1898..2182,2280..2447,2534..>2942)

CDS
 /gene="Me45"
 /product="male fertility protein"
 join(1392..1768,1898..2182,2280..2447,2534..2942)
 /gene="Me45"
 /codon_start=1
 /product="male fertility protein"
 /protein_id="AAK52489.1"
 /db_xref="GI:14028757"
 /translation="MEKRLQMRGRDGIYQYPIHEFPAALALALVDPFGSLPAEVDYRPVGHLELPYGEVWSPRDNASRLRGRLFVGEVGPESIEFDLQGRPYAGLADGRVVMGESEAGMETFAVMPDMSSEVAGVNSTTRQHEKEEFQGLGRFEGETGLVYVADAYYGLMVGOSGVASVAREADGPIRPANDLDVHRGSGVFPDTSRYSRDLNLTLLBEGTGRLLRYDPETSGVAVLKGLVFPVGVQVQISDHOHLLFSETTNCRIKRYVLEBPFRAGEVVPANLGEFPNVANSNGQPVWAIIDCPTPAQGVPRAPRPMIRLTFFKPLSLKVLTMFAARRMHTVLALLDBGRVVEVLSDRGHEVMKLVSEVEGRKLMIGTVAHNHIAITIPYPLD"

ORIGIN
 Query Match 100.0%; Score 157; DB 8; Length 3343;
 Best Local Similarity 100.0%; Pred. No. 1.9e-40;
 Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTTGTCATCGTCC 60
 |||||
 DB 1155 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTTGTCATCGTCC 1214
 |||||

QY 61 AAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATG 120
 |||||
 DB 1215 AAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGATG 1274
 |||||

QY 121 CAACCTTCATGCAAAACGACATATGTTCTCGAAG 157
 |||||
 DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCGAAG 1311
 |||||

RESULT 7
 AX224396 158 bp DNA linear PAT 10-SEP-2001
 LOCUS Sequence 3 from Patent WO0160997.
 ACCESSION AX224396
 VERSION AX224396.1 GI:15554638
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1
REFERENCE Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
AUTHORS Male tissue-preferred regulatory region and method of using same
TITLE Patent: WO 0160997-A 3 23-AUG-2001;
JOURNAL PIONEER HI-BRED INTERNATIONAL, INC. (US)
REFERENCE Location/Qualifiers
AUTHORS 1..158
TITLE /organism="Zea mays"
JOURNAL /mol_type="unassigned DNA"
REFERENCE /db_xref="taxon:4577"

ORIGIN
 Query Match 93.0%; Score 146; DB 6; Length 158;
 Best Local Similarity 99.4%; Pred. No. 7.2e-37;
 Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC 59
 |||||
 DB 1 CGTGCATCTCAGATGGCATCTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC 60
 |||||

QY 60 CAAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGAT 119
 |||||
 DB 61 CAAGCTTGCCCTATTCTGAACCAAGAGATACCTACTCCCAACAATCATCTTACTGAT 120
 |||||

QY 120 GCAACTTCATGCAAAACGACATATGTTCTCGAAG 157
 |||||

Db	121	GCAACTTCATGCAACAGCACATATGTTTCTGAAAC	158
RESULT 8	AC135206	137327 bp	DNA linear
LOCUS	Oryza sativa (japonica cultivar-group)	1linear	PLN 16-APR-2003
DEFINITION	Oryza sativa (japonica cultivar-group)	chromosome 3 clone	
ACCESSION	OJ1041F02	complete sequence.	
VERSION	AC135206		
KEYWORDS	AC135206.3	GI:27596977	
SOURCE	HTG.		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Oryza sativa (japonica cultivar-group)		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.		
REFERENCE	1 (bases 1 to 137327)		
AUTHORS	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K.		
TITLE	Rice Genomic Sequence		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 137327)		
AUTHORS	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saeki,C., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (09-OCT-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	3 (bases 1 to 137327)		
AUTHORS	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saeki,C., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (06-NOV-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	4 (bases 1 to 137327)		
AUTHORS	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (11-JAN-2003) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	5 (bases 1 to 137327)		
AUTHORS	Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T., Saeki,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and Sun,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (16-APR-2003) Clemson University Genomics Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA		
COMMENT	On Jan 11, 2003 this sequence version replaced gi:24635891. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone; and the assembly was confirmed by restriction digest. There are transposons located at area 2947-32874 and at area 19945-33212. Region 70511-70536 is covered by a single subclone. Region 18678-18873 is a double stranded single subclone. Areas 69957-70761, 89836-90416, 98951-99009, 105316-105607, and 133052-133152 are covered by Monsanto only. The nucleotide sequence of this BAC clone was generated by combining Snygenta, Monsanto and Arizona Genomics Institute sequencing data.		
FEATURES	Location/Qualifiers		
source	1. 137327		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39947"		
	/chromosome="3"		
	/clone="OJ1041F02"		
	/notes="(japonica cultivar-group)"		
	/complement(638..9495)		
	/gene="OJ1041F02.1"		
gene	/note="unknown protein"		
CDS	complement(Join(638..838,1141..1276,3133..3244,3326..3421,3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,		
	5322..5411,5484..5569,5808..5942,6035..6538,6622..6837,7548..7593,7679..7815,8196..8328,9221..9495))		
	/gene="OJ1041F02.1"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06844.1"		
	/db_xref="GI:29893590"		
	/translation="MAOSLELLIOFLMPDNDARARAEQIRRLARDPPVVALVHLLRTAKTPNRQLAAVLIRKTKITSHMKLPHPHAKASIKQALIDSTITDHSILVRASANYVSIINAKIVPAGEMPFLPFIQCSQSFQEDHREVALILFSLSTETIGTFOSHILANOPTILKCIQDEKSSRRVRIALKAVGSFLIYVNDGDVVKIFPDVPVSLIYVNSQCLAN GEEVDASIAFELIDELIESPAFLDYSVRSIVQFLEVCNSDELINIRQQAIIQIISM LVKFAESFLKHKHIVIPILQVWCPLLTETADDEGSDLAADSEINIVDTMAINIPRH VPPPLAEASVFRHINPKYBAATATSGVSEGCCHLKDLECLKAVLEALDOE OMVRPAASAPAGQFAPEHLOPLTISYVESVLPCTIINALDPSPEVSEKSYVALAACED MENILPLIDPLMKCLVNSLQSPNNLDCTCSAIGVAAABEQPMYAEVYLEMK GPMVLTNDEDLCARARATEVGIYVMAAGRAMETILPFLTALISGVLDSILEY THGFPSNVAELIDSEFAQYLPHVPLVLAAGSACNLDDGSAVDIDDASVNGSGVSSDD DVNDEPRVRNLSMLTIIADIILTAIRAIIPPAHADVLEKQDILDTVMNIYIKTRBDD KEVVAOACTSLADIYRDCGPALIEBYIRRLADATLILROSCCOQVESDDEGDDID HDEVIMDAVSPDLPAFAVMGSYRPIFTKL PDSIMKPKASKSPHPODCTMYATLAEV AOGMGAPISAYVDKIMPLVLEKIASSEATRNRRNAFCYGEKCRNGAALAKYIGITLH GILHRFAUSEPDDAVRDNAAGAIAMINWQPSIPLNOLPLVPVIVIAFAQVAVSPN ESDEKTVVAKAVASHLISVYGQWQPIISALPFAHNAALASFANRR"		
	12372..15503		
	/note="OJ1041F02.2"		
	/gene="unknown protein"		
	Join(12372..12757,12858..13025,13130..13295,13776..13826,13982..14130,14224..14354,14459..14558,15197..15503)		
	/gene="OJ1041F02.2"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06845.1"		
	/db_xref="GI:29893591"		
	/translation="MPPOGSECRHGWLGVEBAVGGSEPFVYPLRKRLSVDSKAS IAPRICWECDEGADITCDIVAAPLRKSCAKAMPPEAPLFRMTPEPPRQORDGE BARRPGEAIIKRGHSYSIMLQIGISYVGSALPQKLAASDPDEKREKWTFFP EGSKTPPHNSVDPKWDYCPAVPHRLKIGVDAEYVLAIIICGDTLRBLASPKS GCFRTODDRMWTIKVYKSEYKVLIRMSYVENHROYKSTLIRPYCTHCKQKCP KIHRRFDLKGSSHGRTIDTKRKIDETTLKDLQIVAFRIORFYEELMKQIQDCT FLETQGLMDYSLGLVHFRNDYSVKIGISQHIAPKSTGRKSGSSFCGLCFVE SGCXDRDLIDSRKEPIQIGIMPMQAEKSSKILDNPLNRLHFTTPSGGSCDYVL FEGILIDQDYDTIKLEHAYKSFQVNDYISAVDPKLYSRFPQFIRRVIFKEQ"		
	16734..16878		
	/note="putative MITE, Gaijin/Gaigin-like"		
	19263..19454		
	/note="putative transposon, CACTA, En/Spm-like"		
	21819..24660		
	/gene="OJ1041F02.3"		
	/note="hypothetical protein"		
	Join(21819..22007,22096..22206,23772..23996,24523..24660)		
	/gene="OJ1041F02.3"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="AAP06846.1"		
	/db_xref="GI:29893592"		
	/translation="MATASDGGSGGGEORRLISIPKEGERIIAPRRPDGTLRKAIRI RAGVROEVAIYQSKGQMKSGEDVPCYDPAIDAPKTKAARNRREKREKQAS TTNDYKGLHIEDGETDNPDAVDSYTKOISGIAISESLVATSYTDADNDSSES SAPIDDKIRALKKIRLAEAOVGDPENLKPEOLEKMKIKEGWKEELKLLENKSSPA AS"		
	complement(26563..30575)		
	/gene="OJ1041F02.4"		
	/note="unknown protein"		
	complement(Join(26563..26772,27082..27288,27430..27537,27881..27940,28051..28126,28327..28484,28616..28678,28790..28874,29560..30575))		
	/gene="OJ1041F02.4"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06847.1"		
	/db_xref="GI:29893593"		
	/translation="MGPPAAPDPSSSSSGSSRRRRRLRLRRNNAKHIGYDASNFCAYP		

124 CTTCCATGCA 133
| | | | | | | |
119439 CCTTCATGCA 119430

LOCUS	174806 bp	DNA	linear	HTG	29-SEP-2004
DEFINITION	BX950854				
	Danio rerio clone CH211-125M22, WORKING DRAFT SEQUENCE, 5 unordered pieces.				


```

KEYWORDS
SOURCE
ORGANISM
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 174806)
McLaren,S
REFERENCE
AUTHORS
Submitted
JOURNAL
Submitted (28-Sep-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgehire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk
zf1sh-help@sanger.ac.uk
On Sep 25, 2004 this sequence version replaced gt:42821090.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zc125M22
----- Summary Statistics
Assembly program: XGAP; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 172622 bases at least Q40
Consensus quality: 173174 bases at least Q30
Consensus quality: 173177 bases at least Q20
Insert size: 174406; sum-of-contigs
Insert size: 182163; 2.3% error; agarose-gel
Quality coverage: 11.29x in Q20 bases; sum-of-contigs Quality
coverage: 11.03x in Q20 bases; agarose-gel
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 74620: contig of 74620 bp in length
* 74621 74720: gap of 100 bp
* 74721 133368: contig of 58648 bp in length
* 133369 133468: gap of 100 bp
* 133469 150793: contig of 17325 bp in length
* 150794 150893: gap of 100 bp
* 150894 172651: contig of 21758 bp in length
* 172652 172751: gap of 100 bp
* 172752 174806: contig of 2055 bp in length.
FEATURES
source
location/Qualifiers
1..174806
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-125M22"
/clone_lib="CHOR1-211"
1..74620
/note="assembly_fragment:02018
fragment_chain:1
clone_end:SP6
vector_side:left"
74721..133368
/note="assembly_fragment:00796
fragment_chain:1"
133469..150793
/note="assembly_fragment:00037
fragment_chain:1"
150894..172651
/note="assembly_fragment:00366
fragment_chain:1"
172752..174806
/note="assembly_fragment:00028
clone_end:T7
vector_side:right"

```

ORIGIN
 Query Match 22.2%; Score 34.8; DB 2; Length 174806;
 Best Local Similarity 52.8%; Pred. No. 2.7;
 Matches 75; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 15 TGGCATCTACATGCTTGTTCACCGCTTCGTTGTTCATGCTCCAGCCTTGCTAT 74
 DB 162698 TGGGTTGTCTCATCTTCAACCCCAAGCGTGGGTGTAAACACACCCACCTTTAAAGT 1626399
 QY 75 CTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCANGCACAATTCCATGCAA 134
 DB 162638 AAGTAACATTCATCTCAAAAAAGCCCAAAATATTTCATTTAAATCTGAATCTCAGTTAA 162579
 QY 135 ACACGCACATATGTTTCTCGAA 156
 DB 162578 ACTCCACATTTCTCATTTAA 162557
 RESULT 12
 AC104134
 LOCUS AC104134 110127 bp DNA linear PRI 20-MAR-2002
 DEFINITION Homo sapiens BAC clone RP11-525L16 from 2, complete sequence.
 ACCESSION AC104134
 VERSION AC104134.4 GI:19551199
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 110127)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 110127)
 AUTHORS Tomlinson, C., Haakenson, W. and Doeber, A.
 TITLE The sequence of Homo sapiens BAC clone RP11-525L16
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 110127)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (04-DEC-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 4 (bases 1 to 110127)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (15-FEB-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 REFERENCE 5 (bases 1 to 110127)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Mar 20, 2002 this sequence version replaced gi:18677630.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Entry code: MUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiehs@wustl.edu
 ----- Summary Statistics
 Center project name: H_NH0525L16

 NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.
 This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, B., Tatero, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-554H10, 2000 bp overlap; the clone sequenced to the right is RP11-450B9, 2000 bp overlap. Actual start of this clone is at base position 151996 of RP11-554H10.

FEATURES

source

```
location/Qualifiers
1..110127
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="2"
  /map="2"
  /clone="RP11-525L16"
  /clone_id="RPCI-11"
  701..998
    /rpt_family="Alu"
    999..1297
      /rpt_family="Alu"
      1410..1486
        /rpt_family="L1"
        1564..1678
          /rpt_family="Alu"
          1684..1722
            /rpt_family="AT_rich"
            2338..2456
              /rpt_family="MIR"
              4039..4332
                /rpt_family="Alu"
                4412..4532
                  /rpt_family="Alu"
                  4669..4846
                    /rpt_family="MIR"
                    5878..6173
                      /rpt_family="Alu"
                      6385..6584
                        /rpt_family="MIR"
                        6929..7121
                          /rpt_family="MIR"
                          7167..7312
                            /rpt_family="MIR"
                            7953..8202
                              /rpt_family="Alu"
                              8537..8577
                                /rpt_family="L2"
                                8585..8756
                                  /rpt_family="L1"
                                  8768..8888
                                    /rpt_family="Alu"
                                    9005..9219
                                      /rpt_family="L2"
```

```
repeat_region 9272..9838
  /rpt_family="L1"
repeat_region 9849..9976
  /rpt_family="Alu"
repeat_region 10124..10411
  /rpt_family="Alu"
repeat_region 10483..10759
  /rpt_family="L2"
repeat_region 11215..11264
  /rpt_family="(TA)n"
repeat_region 11362..11476
  /rpt_family="Alu"
repeat_region 11478..11776
  /rpt_family="Alu"
repeat_region 12314..12607
  /rpt_family="Alu"
repeat_region 12778..12943
  /rpt_family="MALR"
repeat_region 12954..12979
  /rpt_family="Alu"
repeat_region 14391..14655
  /rpt_family="Alu"
repeat_region 14670..14939
  /rpt_family="MALR"
repeat_region 15435..15449
  /rpt_family="L2"
repeat_region 15450..15730
  /rpt_family="Alu"
repeat_region 15731..15824
  /rpt_family="L2"
repeat_region 15837..15915
  /rpt_family="MALR"
repeat_region 16150..16214
  /rpt_family="MALR"
repeat_region 16224..16382
  /rpt_family="MER1_type"
repeat_region 16533..16747
  /rpt_family="MIR"
repeat_region 16889..17116
  /rpt_family="L2"
repeat_region 17271..17343
  /rpt_family="L1"
repeat_region 17384..17534
  /rpt_family="MIR"
repeat_region 17545..17843
  /rpt_family="Alu"
repeat_region 17995..18056
  /rpt_family="L2"
repeat_region 18057..18074
  /rpt_family="GA-rich"
repeat_region 18075..18379
  /rpt_family="Alu"
repeat_region 18380..18404
  /rpt_family="GA-rich"
repeat_region 18889..19067
  /rpt_family="MER1_type"
misc_feature 19296..19302
  /note="match to EST BE843354 (NID:G10275732)"
repeat_region 19593..19709
  /rpt_family="Alu"
repeat_region 19890..20158
  /rpt_family="Alu"
repeat_region 20225..20698
  /rpt_family="L1"
repeat_region 20737..20769
  /rpt_family="AT_rich"
repeat_region 20887..21186
  /rpt_family="Alu"
misc_feature 21350..21369
  /note="match to EST BG571642 (NID:G13579295)"
repeat_region 21357..21652
  /rpt_family="Alu"
repeat_region 21653..21957
```



```

repeat_region      22328..22636
                    /rpt_family="Alu"
Query Match      22.0%; Score 34.6; DB 9; Length 110127;
Best Local Similarity 54.3%; Pred. No.3.1;
Matches 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 3 TGTGATCTGCATGAGGATCTACATGCTGTGTTACCGCTGCTGTTTCATGTCACAA 62
DB 8763 TGGGATTACGAGGTGAGCCAGTACCTGGCCACCATCTTTCTTCCTCTCTTCCA 87692
QY 63 GCGTTGCTTATTTGACCAAGGAGTACTTATCCCAACATCATCTACTATGCA 122
DB 87693 TCTCCCTTATTTTACCTTTGTTTCATCTACTCAGAACTCATCTCCATACC 87752
QY 123 ACTTCGATG 131
DB 87753 ACCTCCAG 87761

```

```

RESULT 13
AC136896      214946 bp      DNA      linear      PRI 21-MAR-2003
LOCUS      Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
DEFINITION
AC136896
VERSION
AC136896.6 GI:29135656
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE
AUTHORS      Birren, B., Nusbaum, C. and Lander, E.
TITLE        Homo sapiens chromosome 15, clone RP11-86001
JOURNAL
REFERENCES    2 (bases 1 to 214946)
              1 (bases 1 to 214946)

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS

```

```

Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, S., Arachchi, H.M., Batra, N., Basleten, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collimore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Maddison, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mlenge, V., Murphy, T., Naylor, J.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,

```

```

Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 214946)

```

```

TITLE
JOURNAL
COMMENT

```

```

FEATURES
source
1..214946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-86001"
/clone_11b="RP11-11 Human Male BAC"
/complement(1..362)
/rpt_family="HALL"
/complement(823..1431)
/rpt_family="LIMCS"
913..922
/nc="30 qual SNGL region"
/complement(2941..3295)
/rpt_family="MER31B"
/complement(3307..3311)
/nc="30 qual SNGL region"
3438..3476
/nc="single clone coverage"
3546..3586
/rpt_family="GA-rich"
3601..4006
/rpt_family="MT1A1"
complement(4008..4166)
/rpt_family="MIR"
complement(4580..4711)

```



```

repeat_region      /rpt_family="L3"
                    4731..4751
repeat_region      /rpt_family="AT_rich"
                    /complement(4758..4993)
repeat_region      /rpt_family="MIR"
                    complement(5336..5464)
repeat_region      /rpt_family="L3"
                    5935..6412
repeat_region      /rpt_family="MER39"
                    complement(6466..7069)
repeat_region      /rpt_family="MLT1B"
                    complement(8633..8713)
repeat_region      /rpt_family="MIR"
                    8768..9244
repeat_region      /rpt_family="L1MA2"
                    complement(9299..9498)
repeat_region      /rpt_family="MIR"
                    10252..10587
repeat_region      /rpt_family="MLT1A0"
                    12232..12619
repeat_region      /rpt_family="L1MA5"
                    complement(13183..13354)
repeat_region      /rpt_family="MIR"
                    14074..14500
repeat_region      /rpt_family="L1MB3"
                    15070..15201
repeat_region      /rpt_family="FLAM_C"
                    complement(16057..18621)
repeat_region      /rpt_family="L1MB3"
                    complement(18678..19377)
repeat_region      /rpt_family="L1M1"
                    complement(19375..20142)
repeat_region      /rpt_family="L1MB3"
                    22615..22919
repeat_region      /rpt_family="A1USG"
                    complement(25101..25206)
repeat_region      /rpt_family="HAL1"
                    complement(25691..25925)
repeat_region      /rpt_family="HAL1"
                    25996..26033
repeat_region      /rpt_family="CAGA)n"
                    complement(26059..26582)
repeat_region      /rpt_family="MLT2B3"
                    26645..27455
repeat_region      /rpt_family="MER21B"
                    complement(27582..27833)
repeat_region      /rpt_family="HAL1"
                    complement(28185..28759)
repeat_region      /rpt_family="LTR68"
                    complement(29229..29603)
repeat_region      /rpt_family="THE1A"
                    complement(29771..30036)
repeat_region      /rpt_family="A1Ujb"
                    30056..30083
repeat_region      /rpt_family="(TTTA)n"
                    complement(30090..30369)
repeat_region      /rpt_family="A1UJo"
                    31042..32585
repeat_region      /rpt_family="L1M4"
                    32586..32654
repeat_region      /rpt_family="(TCTA)n"
                    32655..33163
repeat_region      /rpt_family="L1M4"
                    33164..33473
repeat_region      /rpt_family="A1Uy"
                    33474..33683
repeat_region      /rpt_family="L1M4"
                    33686..33751
repeat_region      /rpt_family="MLT1D"
                    36941..37015
repeat_region      /rpt_family="MIR"
                    complement(38001..38449)
repeat_region      /rpt_family="L1MD1"

```

```

repeat_region      38453..38855
                    /rpt_family="MLT1D"
repeat_region      complement(38856..39656)

Query Match      21.9%; Score 34.4; DB 9; Length 214946;
Best Local Similarity 54.8%; Pred. No. 3.7;
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY      5 TCATCTCAGTCGATCTACATGCTTGTCAACCGTTCCTTGTTCATGTCGCAAGC 64
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35745 TCATCTCTATGGGACAGAAAAATCATTTGACAAAATATGACATCTTTCATGATGTAAAC 35804
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      65 CTGCGCTATTCTGAACCAAGAGATACCTACTCCCAAAATTCATCTACTATGCAAC 124
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      35805 ATCCAACTACTATGAAGAAAGAGTCTTTCTTAACCAATAAATGTAACCTATGAAAA 35864
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

QY      125 TTCC 128
        ||||
Db      35865 ATCC 35868

RESULT 14
CR847532/c      141810 bp   DNA   linear   HTG 03-OCT-2004
LOCUS
DEFINITION
DANIO rerio clone DKRYP-86C9, *** SEQUENCING IN PROGRESS ***, 8
unordered pieces.
ACCESSION
CR847532
VERSION
CR847532.2 GI:53755907
HTG; HTGS PHASB1.
KEYWORDS
Danio rerio (zebrafish)
SOURCE
Danio rerio
ORGANISM
Buckayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 141810)
McLay, K.
Direct Submission
Submitted (02-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 3, 2004 this sequence version replaced gi:53145936.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK086C9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 13909 bases at least Q40
Consensus quality: 14045 bases at least Q30
Consensus quality: 140735 bases at least Q20
Insert size: 14110; sum-of-contrigs
Insert size: 140600; 2.5% error; agarose-fp
Quality coverage: 8.52x in Q20 bases; sum-of-contrigs Quality
coverage: 8.64x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 19172: contig of 19172 bp in length
* 19173 19272: gap of 100 bp
* 19273 46769: contig of 27497 bp in length
* 46770 46869: gap of 100 bp
* 46870 52415: contig of 5546 bp in length
* 52416 52515: gap of 100 bp

```



```
* 52516 64328: config of 11813 bp in length
* 64329 64428: gap of 100 bp
* 64429 66224: config of 31796 bp in length
* 96325 96324: gap of 100 bp
* 96325 121956: config of 25632 bp in length
* 121957 122056: gap of 100 bp
* 122057 131332: config of 9276 bp in length
* 131333 131432: gap of 100 bp
* 131433 141810: config of 10378 bp in length.

FEATURES
    source
        1..141810
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone="DKEY-86C9"
            /clone_11b="DantioKeyplot"
            1..19172
                /note="assembly: fragment:00300
                fragment chain:1"
            19273..46769
                /note="assembly: fragment:00469
                fragment chain:1"
            46870..52415
                /note="assembly: fragment:00017
                fragment chain:1"
            52516..64328
                /note="assembly: fragment:00167
                fragment chain:1"
            64429..96224
                /note="assembly: fragment:00976
                fragment chain:2"
            96325..121956
                /note="assembly: fragment:01302
                fragment chain:2"
            122057..131332
                /note="assembly: fragment:00056.0"
            131433..141810
                /note="assembly: fragment:00703"

ORIGIN
Query Match      21.5%; Score 33.8; DB 2; Length 141810;
Best Local Similarity 52.5%; Pred. No. 5.8;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY      12 ACATGGCATACTGATCGTGTGTCACCGTGTGCTGTGTCATGTCAGCGCTTGCGCT 71
      |||
DB      41846 ACAGTACCCCTGATCCTTGTGCTGCTCAATGATATTTGACAGTGTACAGAGCTTGGGT 41787

QY      72 ATTCTGAACCAAGAGATACCTACTCCCAAAATCCATCTTACTCATGCAACTTCCATG 131
      |||
DB      41786 AAGTGGGAATATGGGATCATCTGCTCGTGAATAATGATAATATACACAAGAAACTT 41727

QY      132 CAACACGACATATGTTTCC 152
      |||
DB      41726 TCATAAAAAAACTTATTAGTCC 41706

RESULT 15
BX640577      214178 bp      DNA      linear      VRT 10-JUL-2004
LOCUS      Zebrafish DNA sequence from clone DKEY-117014 in linkage group 14,
DEFINITION      complete sequence.
ACCESSION      BX640577
VERSION      BX640577.17 GI:50080105
KEYWORDS
SOURCE      HTG.
ORGANISM      Danio rerio (zebrafish)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE      1 (bases 1 to 214178)
AUTHORS      McGulire,S.
TITLE      Direct Submission
```

```
JOURNAL
Submitted (10-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 8, 2004 this sequence version replaced GI:49659408.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/Clone-derived
Zebrafish PUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhifeng Bao and Sean Eddy, submitted), and those
beginning 'drr' were identified by Rick Waterman (Stephen Johnson
lab, Washu). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-117014
is from a Zebrafish BAC library
VECTOR: pindigOBAC-5.

FEATURES
    source
        1..214178
            /organism="Danio rerio"
            /mol_type="genomic DNA"
            /db_xref="taxon:7955"
            /clone="DKEY-117014"
            /clone_11b="DantioKey"

ORIGIN
Query Match      21.5%; Score 33.8; DB 5; Length 214178;
Best Local Similarity 52.5%; Pred. No. 5.9;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY      12 ACATGGCATACTGATCGTGTGTCACCGTGTGCTGTGTCATGTCAGCGCTTGCGCT 71
      |||
DB      58301 ACAGTACCCCTGATCCTTGTGCTGCTCAATGATATTTGACAGTGTACAGAGCTTGGGT 58360

QY      72 ATTCTGAACCAAGAGATACCTACTCCCAAAATCCATCTTACTCATGCAACTTCCATG 131
      |||
DB      58361 AAGTGGGAATATGGGATCATCTGCTCGTGAATAATGATAATATACACAAGAAACTT 58420

QY      132 CAACACGACATATGTTTCC 152
      |||
DB      58421 TCATAAAAAAACTTATTAGTCC 58441
```

Search completed: September 16, 2005, 03:01:36
Job time : 528.823 secs

CC	21	30	19.1	378	10	ADD33403	ADD33403 Mouse mlt
CC	22	30	19.1	396 <td>10</td> <td>ADD33404</td> <td>ADD33404 Mouse mlt</td>	10	ADD33404	ADD33404 Mouse mlt
C	23	29.8	19.0	373	4	AAH13328	Aah13328 Human CDN
C	24	29.6	18.9	1185	10	ADP00796	Adp00796 Bacterial
C	25	29.4	18.7	110000	3	AAI22303_2	Continuation (3 of
C	26	29.2	18.6	1475	9	AAI62902	Continuation (7 of
C	27	29.2	18.6	110000	2	AAI42063_06	Continuation (7 of
C	28	28.8	18.3	1500	8	ADA68284	Ada68284 Arabidops
C	29	28.6	18.2	756	13	ADSS57339	Adss57339 Bacterial
C	30	28.6	18.2	778	4	AA532892	AA532892 Human gen
C	31	28.6	18.2	778	4	AA532893	AA532893 Human gen
C	32	28.6	18.2	1043	10	ADD171973	Add171973 Human utr
C	33	28.6	18.2	1660	6	ABN67847	Abn67847 Streptoco
C	34	28.6	18.2	5165	11	ACN45165	Acn45165 Human gen
C	35	28.6	18.2	110000	13	ADAB32806_5	Continuation (6 of
C	36	28.6	18.2	164702	8	ACF62730	Acf62730 Cancer ba
C	37	28.6	18.2	164702	8	ADB20845	Adb20845 MRP1 base
C	38	28.6	18.2	164702	10	ADB87934	Adb87934 Human UGT
C	39	28.6	18.2	164702	10	ADB96917	Adb96917 Human MDR
C	40	28.6	18.2	164702	10	ADB92108	Adb92108 Human MDR
C	41	28.4	18.1	1478	3	AAI15843	AAI15843 Human pro
C	42	28.4	18.1	1715	8	ABA00762	Abao00762 P5CR rela
C	43	28.4	18.1	1791	12	ADDO6652	Adog6652 Human tun
C	44	28.4	18.1	1791	12	ADQ85536	Adq85536 Human tun
C	45	28.4	18.1	1791	12	ADQ84654	Adq84654 Human tun

ALIGNMENTS

RESULT 1

AAH76340 standard; DNA; 255 BP.

AAH76340:

AAH76340: 29-OCT-2001 (first entry)

Z. mays Ms45 promoter fragment.

Ms45; male tissue; regulatory region; transcription; male fertility; hybrid seed; promoter; ds.

Zea mays.

WO200160997-A2.

23-AUG-2001.

13-FEB-2001; 2001WO-US004527.

15-FEB-2000; 2000US-00504487.

(PION-) PIONEER HI-BRED INT INC.

Albertsen WC, Fox TW, Garnaat CW, Huffman G, Kendall TL; WPI; 2001-514772/56.

A male tissue-preferred regulatory region comprising nucleotide sequences essential for initiating transcription of the Ms45 gene useful for mediating fertility in a male plant.

Example 5; Fig 8; 50pp; English.

The invention provides a male tissue-preferred regulatory region (I) comprising nucleotide sequences essential for initiating transcription of the Ms45 gene. A method of mediating male fertility in a plant is provided that involves introducing an expression vector comprising a promoter operably linked to (I) into a plant where the exogenous gene impacts male fertility of the plant and (I) controls expression of the exogenous gene. A method of producing hybrid seeds is also provided. The present sequence represents a Z. mays Ms45 promoter fragment


```
XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 157; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 60
DB 15 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 74
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 75 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 134
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 135 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 171

RESULT 2
AA07408 standard; DNA; 1394 BP.
XX AAX07408;
XX 08-JUN-1999 (first entry)
XX DE Zea mays M45 male tissue-preferred regulatory region.
XX KW M45; male; tissue-preferred; regulatory region; plant cells;
XX KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX OS Zea mays.
XX PA WO9859061-A1.
XX PN 30-DEC-1998.
XX PD 19-JUN-1998; 98WO-US012895.
XX PF 23-JUN-1997; 97US-00880499.
XX PR (PION-) PIONEER HI-BRED INT INC.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX DR WPI; 1999-105628/09.
XX CC The sequence is that encoding an M45 male tissue-preferred regulatory
XX CC region. It may be used in the construction of a vector for a method of
XX CC producing exogenous genes in a male tissue-preferred manner, which is
XX CC useful in restoring or conferring fertility, such as in hybrid seed
XX CC production. In conferring fertility, a monocot/dicot plant is transformed
XX CC with the exogenous nucleotide sequence (a male sterility gene, preferably
XX CC M45), which encodes a product selected from auxins, rolb and diptheria
XX CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX CC and infertile plants
XX SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 60
DB 1155 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 1214
```

```
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 3
AA07409 standard; DNA; 1394 BP.
XX AAX07409;
XX 08-JUN-1999 (first entry)
XX DE Zea mays M45 male tissue-preferred regulatory region.
XX KW M45; male; tissue-preferred; regulatory region; plant cells;
XX KM plant tissue; differentiated; hybrid seed; fertility; ss.
XX OS Zea mays.
XX PA WO9859061-A1.
XX PN 30-DEC-1998.
XX PD 19-JUN-1998; 98WO-US012895.
XX PF 23-JUN-1997; 97US-00880499.
XX PR (PION-) PIONEER HI-BRED INT INC.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX DR WPI; 1999-105628/09.
XX CC The sequence is that encoding an M45 male tissue-preferred regulatory
XX CC region. It may be used in the construction of a vector for a method of
XX CC producing exogenous genes in a male tissue-preferred manner, which is
XX CC useful in restoring or conferring fertility, such as in hybrid seed
XX CC production. In conferring fertility, a monocot/dicot plant is transformed
XX CC with the exogenous nucleotide sequence (a male sterility gene, preferably
XX CC M45), which encodes a product selected from auxins, rolb and diptheria
XX CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX CC and infertile plants
XX SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 60
DB 1155 CGTGCATCTCAATGAGCATACATGCTGTTCAACCGTTGCTGTTCCATGCTCC 1214
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTTATTTGAAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311
```


RESULT 4
AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001MO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTGTCATCTCACATGCGCATCTACTAGCTGTGTTCAACGGTTCGTTGTCATGCTCC 60
DB 1155 CGGTGTCATCTCACATGCGCATCTACTAGCTGTGTTCAACGGTTCGTTGTCATGCTCC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGACGACATATGTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGACGACATATGTTCTGTAAC 1311
RESULT 5
AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.

XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001MO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTGTCATCTCACATGCGCATCTACTAGCTGTGTTCAACGGTTCGTTGTCATGCTCC 60
DB 1155 CGGTGTCATCTCACATGCGCATCTACTAGCTGTGTTCAACGGTTCGTTGTCATGCTCC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGACGACATATGTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGACGACATATGTTCTGTAAC 1311
RESULT 6
AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.
PF
XX 15-FEB-2000; 2000US-00504487.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
PS
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
SO

Query Match 93.0%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 5.4e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGGTGATCTCAATGCGATCTACTGCTGTTGTCACCGTTGTC-TTGTTCATCGTC 59
DB 1 CGGTGATCTCAATGCGATCTACTGCTGTTGTCACCGTTGTTGTTCCATCGTC 60
QY 60 CAAGCCTTGCTTATCTGAACCAAGAGATACCTACTCCCAACATTCATTACTCAT 119
DB 61 CAAGCCTTGCTTATCTGAACCAAGAGATACCTACTCCCAACATTCATTACTCAT 120
QY 120 GCAACTTCCATGCAACACGACATATGTTTCTGTAAC 157
DB 121 GCAACTTCCATGCAACACGACATATGTTTCTGTAAC 158

RESULT 7
AAH76336
ID AAH76336 standard; DNA; 50 BP.
XX
XX AAH76336;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX
XX WO200160997-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX
XX 15-FEB-2000; 2000US-00504487.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI

XX WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
PS
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
SO

Query Match 31.8%; Score 50; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 7.2e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 CTGAACCAAGAGATACCTACTCCCAACATTCATTACTGACGAAC 124
DB 1 CTGAACCAAGAGATACCTACTCCCAACATTCATTACTGACGAAC 50

RESULT 8
AAH76337
ID AAH76337 standard; DNA; 40 BP.
XX
XX AAH76337;
AC
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
OS
XX
XX WO200160997-A2.
PN
XX
XX 23-AUG-2001.
PD
XX
XX 13-FEB-2001; 2001WO-US004527.
PF
XX
XX 15-FEB-2000; 2000US-00504487.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI WPI; 2001-514772/56.
DR
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the MS45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 14; Page 32; 50pp; English.
PS
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC Present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays M45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 AGGATACCTACTCTCCCAACAATCCATCTTACTCTATGCAAC 124
Db 1 AGGATACCTACTCTCCCAACAATCCATCTTACTCTATGCAAC 40
RESULT 9
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KM Mouse; 85; EST; atopic dermatitis; psoriasis; dermatological;
XX anti-inflammatory; antipsoriatic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN MO2004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003WC-JP009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
PR 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
PA (UYJU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
PI Mitsutshiri K;
XX
DR WPI; 2004-214514/20.
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
PT expression of an indicator gene at a rash site and non-rash site of a
PT person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484bp; Japanese.
XX
XX The invention relates to detecting atopic dermatitis or psoriasis
CC comprising assaying the levels of expression of an indicator gene at a
CC rash site and non-rash site of a person with atopic dermatitis or
CC psoriasis, comparing these levels with those of a healthy person, and
CC determining that if the levels of indicators are higher or lower, then
CC this indicates the disease. Also included are a reagent for detecting
CC atopic dermatitis or psoriasis, a kit for screening for treatments, a
CC transgenic non human vertebrate animal models for the diseases, an agent
CC for inducing the diseases in mice and a DNA chip for assaying for the
CC indicator genes. The method is used for treatment, detection and animal
CC models for research of atopic dermatitis and psoriasis. The present
CC sequence is a Mouse atopic-dermatitis/psoriasis-associated EST
CC (expressed sequence tag).
XX
SQ Sequence 3267 BP; 1029 A; 658 C; 590 G; 990 T; 0 U; 0 Other;
Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.2;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 27 TGCTTGTCACACGCTTGCTTGTCATGTCACAGCTTGCTTATTCGAACCAAGAG 86

Db 2774 TTCTTTTTCACACTCTCCTTGATTTCTTTGTAACCTCACTTCATTTTGCCCTTTAAGTT 2833
QY 87 GATACCTACTCCCAACAATTCATCTTACTCTGACCACTTCGCAACACGCACATAT 146
Db 2834 GTTCTGATTTCCCAAGAAAGCCATTTTGCACTGACGTCAAAAGATTAACACCTGAAAT 2893
QY 147 GT 148
Db 2894 GT 2895
RESULT 10
ACA48402/C
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Streptococcus mutans.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WC-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELITR-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyckind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
PT P-PSDB; ABU44532.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation or
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 14; SEQ ID NO 36272; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2352 BP; 762 A; 401 C; 500 G; 689 T; 0 U; 0 Other;

XX Query Match 20.1%; Score 31.6; DB 8; Length 2352;

XX Best Local Similarity 58.5%; Pred. No. 3.8;

XX Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

XX Db 62 AGCCTTGCCCTATTCGAAACCAAGATACCTACTCCCAACATCTTACTGATGC 121

XX 234 ATCCTGACTGTGCCCAAGCAAGAGATACCTCTGAAAATCTGTCTACTGCTGC 175

XX Qy 122 AACTTCATGCAACACGACGACATATGTTCTCTGA 155

XX Db 174 AATTTCAGCAAAAGAACGTTCAATTTTCTCTGA 141

XX RESULT 11

XX ABN80329/c

XX ID ABN80329 standard; DNA; 15734 BP.

XX AC ABN80329;

XX DT 15-JUL-2002 (first entry)

XX DE Human chemically modified disease associated gene SEQ ID NO 346.

XX KM Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;

XX KM heart disease; epilepsy; histone deacetylation; muscular dystrophy;

XX KM dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;

XX KM antiabiotic; cytostatic; anticonvulsant; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200200927-A2.

XX PD 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP007536.

XX PR 30-JUN-2000; 2000DE-01032529.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PA (EPig-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX PI WPI; 2002-130908/17.

XX DR Novel nucleic acid useful for diagnosis and therapy of diseases

XX PT associated with development genes such as diabetes, comprises a sequence

XX PT of a segment of chemically pretreated DNA of genes associated with

XX PT development.

XX Claim 1; SEQ ID NO 346; 27pp; English.

XX The invention relates to a nucleic acid (I) comprising a sequence at

XX least 18 bases in length of a segment of chemically pretreated DNA (II)

XX of genes associated with development selected from 87 genes listed in the

XX specification such as ACCPN, ADPN, or AFDI and comprising one of 350

XX sequences (ABN79984-ABN80333) or their complements. The invention is

XX useful for the diagnosis or therapy of diseases associated with

CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Curriano syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs). Note: The sequence data for this
CC patent did not form part of the printed specification but is based on
CC sequence information supplied to Derwent by the European Patent Office

XX Sequence 15734 BP; 3630 A; 744 C; 4234 G; 7126 T; 0 U; 0 Other;

XX Query Match 20.0%; Score 31.4; DB 6; Length 15734;

XX Best Local Similarity 59.6%; Pred. No. 8;

XX Matches 53; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

XX Qy 52 CGATGTCACACCTTGCTATTCGAAACCAAGATACCTTCCCAACATCCATC 111

XX Db 8802 CCTCTTCTCTCCCTTATCGAATCTTAAACAATCAAAACCTTAAATTAACACCAAC 8743

XX Qy 112 TTAATGACCACTTCATGCAACACGC 140

XX Db 8742 TACCTAACCCATCTTCCAAATTAACACC 8714

XX RESULT 12

XX AAF22279

XX ID AAF22279 standard; DNA; 64415 BP.

XX AC AAF22279;

XX DT 20-MAR-2001 (first entry)

XX DE BAC containing repeats from centromeres 1-4 #2.

XX KM Centromere; mitochondria; vector; ds.

XX KM Arabidopsis thaliana.

XX OS Arabidopsis thaliana.

XX OS WO200055325-A2.

XX PN 21-SEP-2000.

XX PD 17-MAR-2000; 2000WO-US007392.

XX PF 18-MAR-1999; 99US-0125219P.

XX PR 01-APR-1999; 99US-0127409P.

XX PR 18-MAY-1999; 99US-0134770P.

XX PR 13-SEP-1999; 99US-0153584P.

XX PR 17-SEP-1999; 99US-0154603P.

XX PR 16-DEC-1999; 99US-0172493P.

XX PA (UYCH-) UNIV CHICAGO.

XX PI Preuss D, Copenhaver G, Keith K;

XX PI WPI; 2000-587529/55.

XX DR Recombinant DNA construct comprising a plant centromere, useful for

XX PT producing stably inherited mitochondria which can serve as vectors for the

XX PT construction of transgenic plant and animal cells.

XX Claim 102; Page 321-335; 1449pp; English.

XX The present invention relates to a recombinant DNA construct of a plant

XX (Arabidopsis thaliana) centromere. The constructs are useful for

XX producing stably inherited mitochondria which can serve as vectors for the

XX construction of transgenic plant and animal cells expressing selected

XX proteins such as hormones, enzymes, interleukins, clotting factors,

CC cytokines, antibodies, and growth factors
XX
SQ Sequence 64415 BP, 18698 A, 13554 C, 13083 G, 18980 T, 0 U, 100 Other;
Query Match 19.5%; Score 30.6; DB 3; Length 64415;
Best Local Similarity 58.1%; Pred. No. 23;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 65 CTGGCCATTCTGACCAAGAGTACTCTCCCAACCAATCCATTCTCTATGCAAC 124
DB 7117 CTGGCTCTTTAGATATTAATGATATTTCTCTCAACCACTAAACCTTAACCTTACACC 7176
QY 125 TTCCATGCAACACGACATATGTTTCTCTGAC 157
DB 7177 TTAATCCAAACCTTAATCTAATCTCTTAAC 7209
RESULT 13
ADA02717
ID ADA02717 standard; DNA: 79467 BP.
XX
AC ADA02717;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
KM Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukemia; cytostatic; gene therapy; drug screening;
XX gene; ds.
XX
OS Mus sp.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR MPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX
PS Claim 1; SEQ ID NO 1235; 245pp; English.
XX
CC The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-AD03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC leukemia (especially breast cancer, prostate cancer, lymphoma or
CC leukemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 79467 BP, 18256 A, 19684 C, 20062 G, 20204 T, 0 U, 1261 Other;
Query Match 19.5%; Score 30.6; DB 9; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGTCGCAAGCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACATCCAT 110
DB 53779 TCCATGTCGATGATCTTGCTTCTCTGGAACCAAGCTGAGCTTATGACACACACACA 53838
QY 111 CTTACTGATGCAACTTCGATGCAACACGACATATGCTTC 151
DB 53839 CACACACACACACACACACACACACACACACAGCTTGCTC 53879
RESULT 14
ADB72455
ID ADB72455 standard; DNA: 79467 BP.
XX
AC ADB72455;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Nfatc1 gene.
XX
KM mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
XX
OS Mus sp.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
XX
PR 23-OCT-2001; 2001US-00004113.
XX
PR 08-NOV-2001; 2001US-00052482.
XX
PR 30-NOV-2001; 2001US-00997722.
XX
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR MPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
XX cancers, neoplasm, adenocarcinoma, or sarcomas.
XX
PS Claim 1; SEQ ID NO 283; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.
XX
SQ Sequence 79467 BP, 18256 A, 19684 C, 20062 G, 20204 T, 0 U, 1261 Other;
XX
Query Match 19.5%; Score 30.6; DB 10; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGTCGCAAGCTTGCTTATTTGAAACCAAGAGATACCTATCCCAACATCCAT 110
DB 53779 TCCATGTCGATGATCTTGCTTCTCTGGAACCAAGCTGAGCTTATGACACACACACA 53838

QY 111 CTTACTCATGCAACTTTCATGCAACGACGACATATGTTTC 151
 Db 53839 CACACACACACACACACACCAACACACAGCTTGTGGCTC 53879

RESULT 15

AD895965
 ID ADE95965 standard; DNA; 79467 BP.

AC ADE95965;

DT 12-FEB-2004 (first entry)

DE Mouse Nfatc1 gene genomic DNA sequence.

KM cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
 KM lymphoma; breast cancer; prostate cancer; leukaemia; ds; mouse; murine;
 KM Nfatc1.

OS Mus sp.

PN W02003039484-A2.

PD 15-MAY-2003.

PF 08-NOV-2002; 2002MO-US036071.

PR 08-NOV-2001; 2001US-00052482.

PA (SAGR-) SAGRES DISCOVERY.

PI Morris DW, Engelhard EK;

DR WPI; 2003-441462/41.

PT New carcinoma associated nucleic acids and proteins, useful for screening
 PT drug candidates, or for diagnosing and treating carcinomas, e.g.
 PT lymphoma, breast cancer, prostate cancer or leukemia.

PS Claim 1; SEQ ID NO 223; 793pp; English.

CC This invention relates to novel recombinant nucleic acids for use in
 CC diagnosis and treatment of cancer, especially carcinomas, as well as the
 CC use of compositions in screening methods. The compositions of the
 CC invention may have cytostatic activity whilst the disclosed sequences may
 CC be useful for gene therapy. The carcinoma associated nucleic acids and
 CC proteins are useful for diagnosing and treating carcinomas, for example
 CC lymphoma, breast cancer, prostate cancer or leukemia, or for screening
 CC drug candidates or bioactive agents capable of binding to, or modulating
 CC the activity of, a carcinoma associated protein. The present sequence is
 CC the genomic DNA sequence of the mouse Nfatc1 gene which is a carcinoma
 CC associated gene of the invention.

SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

Query Match 19.5%; Score 30.6; DB 10; Length 79467;
 Best Local Similarity 56.4%; Pred. No. 25;

Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCGAAGCCTTGCTATTTGAAACCAAGAGATACCTACTCCCAACATCCAT 110
 Db 53779 TCCACTACTGATCCTTGCTCTCTGGAACAAAGCTGAGCTTAGTCACACACACACA 53838

QY 111 CTTACTCATGCAACTTTCATGCAACGACGACATATGTTTC 151
 Db 53839 CACACACACACACACACACCAACACACAGCTTGTGGCTC 53879

Search completed: September 15, 2005, 21:33:21
 Job time : 77.0101 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 102.248 Seconds

(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1155_1311

Perfect score: 157

Sequence: 1 cgtgcattctacatgcat.....cgcatatgtcttcgaac 157

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues 14778644

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*

22: /cgn2_6/ptodata/2/pubpna/US10I_NEW_PUB.seq:*

23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*

24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	20	US-10-713-381-9
2	157	100.0	1394	20	US-10-713-381-1
3	157	100.0	1394	20	US-10-713-381-2
4	146	93.0	158	20	US-10-713-381-3
5	50	31.8	50	20	US-10-713-381-5
6	40	25.5	40	20	US-10-713-381-6
7	32.2	20.5	399	19	US-10-674-124A-9081

C	8	31.6	20.1	2352	17	US-10-282-122A-16272	Sequence 36272, A
C	9	30.8	19.6	1122	19	US-10-437-963-1496	Sequence 1496, Ap
	10	30.6	19.5	559	13	US-10-027-632-244811	Sequence 244811,
	11	30.6	19.5	559	17	US-10-027-632-244811	Sequence 244811,
	12	30.6	19.5	79467	18	US-10-052-482-223	Sequence 223, App
	13	30.4	19.4	52001	19	US-10-317-273-11	Sequence 11, Appl
	14	30	19.1	30	20	US-10-713-381-4	Sequence 4, Appl
	15	30	19.1	843	13	US-10-027-632-170626	Sequence 170626,
	16	30	19.1	843	17	US-10-027-632-170626	Sequence 170626,
	17	29.8	19.0	236	9	US-09-783-590-10490	Sequence 57018, A
C	18	29.8	19.0	691	18	US-10-424-559-57018	Sequence 57018, A
	19	29.8	19.0	2735	13	US-10-027-632-254612	Sequence 254612,
	20	29.8	19.0	2735	17	US-10-027-632-254612	Sequence 254612,
	21	29.8	19.0	1223197	13	US-10-027-632-179264	Sequence 179264,
	22	29.8	19.0	1223197	17	US-10-027-632-179264	Sequence 179264,
C	23	29.6	18.9	680	19	US-10-767-701-2154	Sequence 2154, Ap
C	24	29.2	18.6	1475	17	US-10-321-802-11	Sequence 11, Appl
C	25	29.2	18.6	1830121	17	US-10-329-670-1	Sequence 1, Appl
C	26	29.2	18.6	1830121	20	US-10-158-865-1	Sequence 1, Appl
C	27	29.2	18.6	1830121	22	US-10-981-687-1	Sequence 78606, A
	28	29	18.5	358	20	US-10-425-115-78606	Sequence 223444,
	29	29	18.5	553	13	US-10-027-632-222444	Sequence 223444,
	30	29	18.5	553	17	US-10-027-632-222444	Sequence 223444,
	31	29	18.5	553	17	US-10-027-632-222444	Sequence 223444,
C	32	29	18.5	553	17	US-10-027-632-222445	Sequence 85365, A
C	33	29	18.5	7192	19	US-10-437-963-85365	Sequence 21721, A
	34	28.8	18.3	732	13	US-10-027-632-21721	Sequence 21721, A
	35	28.8	18.3	732	17	US-10-027-632-21721	Sequence 21721, A
	36	28.8	18.3	732	17	US-10-027-632-21721	Sequence 21721, A
	37	28.8	18.3	732	17	US-10-027-632-21721	Sequence 127179, A
	38	28.8	18.3	733	13	US-10-027-632-127179	Sequence 127179,
	39	28.8	18.3	733	13	US-10-027-632-127180	Sequence 127180,
	40	28.8	18.3	733	17	US-10-027-632-127180	Sequence 53712, A
C	41	28.8	18.3	715517	13	US-10-027-632-53712	Sequence 53712, A
C	42	28.8	18.3	715517	17	US-10-027-632-53712	Sequence 53712, A
C	43	28.8	18.3	715517	20	US-10-425-115-141945	Sequence 141945,
	44	28.6	18.2	220	20	US-10-425-115-17846	Sequence 17846, A
	45	28.6	18.2	280	20	US-10-425-115-17846	Sequence 17846, A

ALIGNMENTS

RESULT 1

US-10-713-381-9

Sequence 9, Application US/10713381

Publication No. US2004022131A1

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, MARC C.

APPLICANT: FOX, TIMOTHY W.

APPLICANT: GARNAT, CARL W.

APPLICANT: HUFFMAN, GARY

APPLICANT: KENDALL, TIMMY L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF

TITLE OF INVENTION: USING SAME

FILE REFERENCE: 578R

CURRENT APPLICATION NUMBER: US/10/713,381

CURRENT FILING DATE: 2003-11-14

PRIOR APPLICATION NUMBER: 08/880,499

PRIOR FILING DATE: 1997-06-23

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patentlin Ver. 2.0

SEQ ID NO 9

LENGTH: 255

TYPE: DNA

ORGANISM: Zea mays

US-10-713-381-9

Query Match 100.0%; Score 157; DB 20; Length 255;

Best Local Similarity 100.0%; Pred. No. 1.7e-43;

Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 cgtgcattctacatgcatctactacatgcttttcacacgcttgcgttttcacatgcttc


```
Db      15 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 74
Qy      61 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 120
Db      75 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 134
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
Db      135 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 171
```

RESULT 2

US-10-713-381-1

; Sequence 1, Application US/10713381

; Publication No. US20040221331A1

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, MARC C.

; APPLICANT: FOX, TIMOTHY W.

; APPLICANT: GARNAAT, CARL W.

; APPLICANT: HUFEMAN, GARY

; APPLICANT: KENDALL, TIMMY L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF

; FILE REFERENCE: 578R

; CURRENT APPLICATION NUMBER: US/10/713,381

; PRIOR FILING DATE: 2003-11-14

; PRIOR APPLICATION NUMBER: 08/880,499

; PRIOR FILING DATE: 1997-06-23

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1394

; TYPE: DNA

; ORGANISM: Zea mays

US-10-713-381-1

Query Match 100.0%; Score 157; DB 20; Length 1394;

Best Local Similarity 100.0%; Pred. No. 3.1e-43;

Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 60
Db      1155 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 1214
```

```
Qy      61 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 120
Db      1215 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 1274
```

```
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311
```

RESULT 3

US-10-713-381-2

; Sequence 2, Application US/10713381

; Publication No. US20040221331A1

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, MARC C.

; APPLICANT: FOX, TIMOTHY W.

; APPLICANT: GARNAAT, CARL W.

; APPLICANT: HUFEMAN, GARY

; APPLICANT: KENDALL, TIMMY L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF

; FILE REFERENCE: 578R

; CURRENT APPLICATION NUMBER: US/10/713,381

; PRIOR FILING DATE: 2003-11-14

; PRIOR APPLICATION NUMBER: 08/880,499

; PRIOR FILING DATE: 1997-06-23

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2
```

Query Match 100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 60
Db      1155 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 1214
Qy      61 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 120
Db      1215 AAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCATG 1274
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTGTAAC 1311
```

RESULT 4

US-10-713-381-3

; Sequence 3, Application US/10713381

; Publication No. US20040221331A1

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, MARC C.

; APPLICANT: FOX, TIMOTHY W.

; APPLICANT: GARNAAT, CARL W.

; APPLICANT: HUFEMAN, GARY

; APPLICANT: KENDALL, TIMMY L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF

; FILE REFERENCE: 578R

; CURRENT APPLICATION NUMBER: US/10/713,381

; PRIOR FILING DATE: 2003-11-14

; PRIOR APPLICATION NUMBER: 08/880,499

; PRIOR FILING DATE: 1997-06-23

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 158

; TYPE: DNA

; ORGANISM: Zea mays

US-10-713-381-3

Query Match 93.0%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 8.2e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy      1 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 59
Db      1 CGGTGCATCTCACATGGCATACTACATGCTGTGTTCAACGGTTCGTCTTGTTCATGCTCC 60
Qy      60 CAAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAT 119
Db      61 CAAGCCTTGCTTATCTTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAT 120
Qy      120 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 157
Db      121 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 158
```

RESULT 5

US-10-713-381-5

; Sequence 5, Application US/10713381

; Publication No. US20040221331A1

; GENERAL INFORMATION:

; APPLICANT: ALBERTSEN, MARC C.

; APPLICANT: FOX, TIMOTHY W.

; APPLICANT: GARNAAT, CARL W.


```

; APPLICANT: HUFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match          31.8%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
75 CTGAACCAAGAGATGATCTCTCCCAACATCTTACTGATGCAAC 124
1 CTGAACCAAGAGATGATCTCTCTCCCAACATCTTACTGATGCAAC 50

Db
1 CTGAACCAAGAGATGATCTCTCTCCCAACATCTTACTGATGCAAC 50

RESULT 6
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match          25.5%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY
85 AGGATACCTACTCTCCCAACATCTTACTGATGCAAC 124
1 AGGATACCTACTCTCCCAACATCTTACTGATGCAAC 40

Db
1 AGGATACCTACTCTCCCAACATCTTACTGATGCAAC 40

RESULT 7
US-10-674-124A-9081/c
; Sequence 9081, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideooh1
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
```

```

; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 9081
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr5.fa.07f1r.139538206
; FEATURE:
; OTHER INFORMATION: Located on chromosome 5
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 128437212
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

Query Match          20.5%; Score 32.2; DB 19; Length 399;
Best Local Similarity 61.2%; Pred. No. 1.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY
72 ATTGACCAAGAGATGATCTCTCCCAACATCTTACTGATGCAACTTCAG 131
239 ATTGACCAAGAGATGATCTCTCTCCCAACATCTTACTGATGCAACTTCAG 180

Db
132 CAAACGCAACATCTTCTCTGAA 156
179 AAAAAACCAACATCTTGTGATGAA 155

RESULT 8
US-10-282-122A-36272/c
; Sequence 36272, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Lianggu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```



```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 36272
/ LENGTH: 2352
/ TYPE: DNA
/ ORGANISM: Streptococcus mutans
US-10-282-122A-36272
```

```
Query Match          20.1%; Score 31.6; DB 17; Length 2352;
Best Local Similarity 58.5%; Pred. No. 3.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 62 AGCCTTGCTATTCTGAACCAAGATACCTACCCAAACATCATCTTACTCATGC 121
    |||||
DB 234 ATCTGACTGTGCCCAAGCCAAAGCAAGACATCTCTGAAAATCTGTGCTACATCTGC 175
    |||||
QY 122 AACTTCATGCAACAACGACATATGTTCTCTGA 155
    |||||
DB 174 AATTTCAGCAAAAAGAAAGCTTCATTCTTCTCTGA 141
    |||||
```

RESULT 9

```
US-10-437-963-1496/C
/ Sequence 1496, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 1496
/ LENGTH: 1122
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496
```

```
Query Match          19.6%; Score 30.8; DB 19; Length 1122;
Best Local Similarity 57.1%; Pred. No. 5.6;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 60 CAAGCTTGCTATTCTGAACCAAGATACCTACCCAAACATCATCTTACTCAT 119
    |||||
DB 313 CAAGCTTCGATATTTCATTCCTCCGGGATCATTAACTCAACAAAGCCATATTCATTCT 254
    |||||
QY 120 GCAACTTCATGCAACAACGACATATGTTCTCTGAAC 157
    |||||
DB 253 TCAGATCCAAATCGACCTCCCATATCTGTCTGAAC 216
    |||||
```

```
RESULT 10
US-10-027-632-244811
```

```
/ Sequence 244811, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ NUMBER OF SEQ ID NOS: 325720
/ SEQ ID NO 244811
/ LENGTH: 559
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-244811
```

```
Query Match          19.5%; Score 30.6; DB 13; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 24 ACATGCTGTTCAACCGTTCGTTGTTCCATGTCGAAGCCTTGCTATTCGAACCA 83
    |||||
DB 380 ACATGATTTTCATGATTTTCGATTTGCTCATCTTGCGAGGTACTTTTCTAGATTCA 439
    |||||
QY 84 GAGGATACCTACTCCCAACCAATCATCTTACTCATGCAACTTCATGCAACAGGCACA 143
    |||||
DB 440 TCTGCTCCCTTACTTTTAAAGATTTCATGGAAGTCTTACAAATCCATAGCACACTGAAC 499
    |||||
QY 144 TATGTTCTCTGA 156
    |||||
DB 500 ATTTTCTTCCCA 512
    |||||
```

RESULT 11

```
US-10-027-632-244811
/ Sequence 244811, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
```



```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244811
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-244811

Query Match      19.5%; Score 30.6; DB 17; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 ACATGCTGTTCACCGTGTCTTTCATGCTGCAAGCTTGCTATTTCGAACCA 83
DB 380 ACATGCTGTTCACCGTGTCTTTCATGCTGCTGAGGACTTTTCTAGATTCA 439
QY 84 GAGGATACCTACTCCCAACATCCATCTTACTGATCAACTTCGACCAACGACGA 143
DB 440 TCGCTCCCTACCTTTAAAGATTCCATGGAAGTCTTCAAAATCCCATACGACCTGAC 499
QY 144 TATGTTTCTGAA 156
DB 500 ATTTCTTCCCA 512

RESULT 12
US-10-052-482-223
; Sequence 223, Application US/10052482
; Publication No. US2004072264E1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 79467
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5502)..(5521)
; OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4095)..(4369)
; OTHER INFORMATION: "n" at positions 4099 to 4369 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10089)..(10620)
; OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13273)..(13370)
; OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (20762)..(20781)
; OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30751)..(30916)
; OTHER INFORMATION: "n" at positions 30751 to 30916 can be any base
; NAME/KEY: misc feature
; LOCATION: (46579)..(46772)
```

```

; OTHER INFORMATION: "n" at positions 46759 to 46772 can be any base
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (77382)..(77401)
; OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

Query Match      19.5%; Score 30.6; DB 18; Length 79467;
Best Local Similarity 56.4%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCAACGCTTGCTATTTCGAACCAAGAGATACCTACCTCCCAACATCCAT 110
DB 53779 TCCATGCTCAACGCTTGCTATTTCGAACCAAGAGATACCTACCTCCCAACATCCAT 53838
QY 111 CTACTCATGCAACTTCATGCAACGACGACATATATTTTC 151
DB 53839 CACACACACACACACACACACACACACACACACAGCTTGCGTC 53879

RESULT 13
US-10-317-273-11
; Sequence 11, Application US/10317273
; Publication No. US20040110158A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
; FILE REFERENCE: RTS-0478
; CURRENT APPLICATION NUMBER: US/10/317,273
; CURRENT FILING DATE: 2002-12-10
; NUMBER OF SEQ ID NOS: 159
; SEQ ID NO 11
; LENGTH: 52001
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
; LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
; LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
; LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
; LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
; LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
; LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
; LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
; LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
; LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 11220
; OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

Query Match      19.4%; Score 30.4; DB 19; Length 52001;
Best Local Similarity 61.2%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 50 TTCATGCTCAACGCTTGCTATTTCGAACCAAGAGATACCTACCTCCCAACATCCAT 109
DB 20055 TTCCTTTTCTAAACTGAAATTCATGCAAAAAAGAGATATCAATTTTAAACTACCA 20114
QY 110 TCTTACTCATGCAACTTCA 129
```


Db 20115 TCTTCTCAGCAACTTCTA 20134

RESULT 14
US-10-713-381-4

Sequence 4, Application US/10713381
Publication No. US20040221331A1
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, MARC C.
APPLICANT: FOX, TIMOTHY W.
APPLICANT: GARNATT, CARL W.
APPLICANT: HUFEMAN, GARY
APPLICANT: KENDALL, TIMMY L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
FILE REFERENCE: 578R
CURRENT APPLICATION NUMBER: US/10/713,381
CURRENT FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 08/880,499
PRIOR FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 30
TYPE: DNA
ORGANISM: Zea mays
US-10-713-381-4

Query Match 19.1%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 CATGCTGTTCACCGTTCGTTGTTCCA 54
DB 1 CATGCTGTTCACCGTTCGTTGTTCCA 30

RESULT 15

US-10-027-632-170626
Sequence 170626, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: PaSeq for Windows Version 4.0
SEQ ID NO 170626
LENGTH: 843
TYPE: DNA
ORGANISM: Human
US-10-027-632-170626

Query Match 19.1%; Score 30; DB 13; Length 843;
Best Local Similarity 51.5%; Pred. No. 9.5;
Matches 69; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 6 CATTCAGATGAGCATATCATGCTGTTCACCGTTCGTTGTTCCATGTCGAAGCC 65
DB 552 CAGCACCCAGGCCAGCTCAAGACAACGTCAGCTCTCTGAGTGGGCCACCATCCATGCC 611
QY 66 TTGCTTATTTCTGAACCAAGAGATACCTATCCCAACATTCATCTTACTGATGCACT 125
DB 612 TTGCATGTGCTGGAAGTGGGGCTTCGCGCTCCCTGCTCATCATCAACGCTGGGAGGCGCTCC 671
QY 126 TCCATGCAAAACACG 139
DB 672 TGCATCCGAACACG 685

Search completed: September 15, 2005, 20:45:58
Job time : 106.248 secs

QY 61 AAGCTTGCTATCTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 120
DB 361 AAGCTTGCTATCTTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 302
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 301 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 265

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS OGIAG08TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA0716B15,
DEFINITION genomic survey sequence.
ACCESSION CG224225 GI:34124113
VERSION CG224225.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 915)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGIAG08TV
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TF
Classes: sheared ends.

FEATURES
source location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0716B15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 Kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 157; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 1,9e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATCTACATGCGATACATACATGCTTGTCAACCGCTTGCTTGCATGTC 60
DB 534 CGTGCATCTACATGCGATACATACATGCTTGTCAACCGCTTGCTTGCATGTC 593
QY 61 AAGCTTGCTATCTTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 120
DB 594 AAGCTTGCTATCTTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 653
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 654 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 690

RESULT 3
CC656933 963 bp DNA linear GSS 19-JUN-2003
LOCUS OGMD020TM_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBMA0554D15,
DEFINITION genomic survey sequence.
ACCESSION CC656933

VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGMD020TV
Contact: Cathy WhiteJaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES
source location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0554D15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 Kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 157; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 2e-39;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGCATCTACATGCGATACATACATGCTTGTCAACCGCTTGCTTGCATGTC 60
DB 752 CGTGCATCTACATGCGATACATACATGCTTGTCAACCGCTTGCTTGCATGTC 811
QY 61 AAGCTTGCTATCTTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 120
DB 812 AAGCTTGCTATCTTGAACCAAGAGATACCTACTCCAAACATCATCTTACTCATG 871
QY 121 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 157
DB 872 CAACCTTCATGCAAAACGACATATGTTCTCTGAAC 908

RESULT 4
CE588104 636 bp DNA linear GSS 28-SEP-2003
LOCUS CE588104/c tigr-gss-dog-1700036359543_Dog_library_Canis_familiaris_genomic,
DEFINITION genomic survey sequence.
ACCESSION CE588104
VERSION CE588104.1 GI:36904885
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Flapsipedia; Canidae; Canis.
1 (bases 1 to 636)
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,W., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
JOURNAL PUBLISHED 22875432
COMMENT 14512627
Contact: Kirkness EF
The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkes@tigr.org
Class: shotgun.

FEATURES

source

1.636
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog library"
/note="Site 1: Bctx; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 24.1%; Score 37.8; DB 9; Length 636;
Best Local Similarity 58.4%; Pred. No. 0.39; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 TCTTGTTCATCGTCCAGCCTTGCTTATCTGAACCAAGAGATACCTAATCCCAAC 104
DB 292 TATTTTCCAGTTCAGCATCTGTTGGACCAAGGTATACCTCATTTTAAAA 233

QY 105 ATCCATCTTACTGATGCACTTCATGCAACACGACATATGTTCTGAAC 157
DB 232 ATACATTTTATATAGAAATTTTCAACACAAAAATAAATGATGATGAC 180

RESULT 5
LOCUS CO689495 643 bp mRNA linear EST 26-JUL-2004
DEFINITION DG11-25h6 DG11-kidney Canis familiaris cDNA 3', mRNA sequence.
ACCESSION CO689495
VERSION CO689495.1 GI:50638161
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 643)
Schlueter, J., Hermann, J., Weindel, M., Schuette, D., Kranz, H.,
Henrich, J., and Loebert, R.
Dog arraying cDNA clone collection
Unpublished (2004)
Contact: Thomas Schlueter
LION bioscience AG
Waldfosserstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schlueter@lionbioscience.com.
Location/Qualifiers
1.643
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/clone_lib="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="DG11-kidney"
/note="Organ: Kidney; Vector: Dog pBluescript LION"

FEATURES
source

ORIGIN

Query Match 24.1%; Score 37.8; DB 7; Length 643;
Best Local Similarity 58.4%; Pred. No. 0.4; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 TCTTGTTCATCGTCCAGCCTTGCTTATCTGAACCAAGAGATACCTAATCCCAAC 104
DB 333 TATTTTCCAGTTCAGCATCTGTTGGACCAAGGTATACCTCATTTTAAAA 274

QY 105 ATCCATCTTACTGATGCACTTCAGCAACGACGATATGTTCCGAC 157
DB 273 ATACATTTTATATAGAAATTTTCAACACAAAAATAAATGATGATGAC 221

RESULT 6
LOCUS AG366773/ 1085 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMG01-170B12.TU, genomic survey sequence.
ACCESSION AG366773
VERSION AG366773.1 GI:47977978
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
BAC end Sequences of Library MSMG01
Unpublished
2 (bases 1 to 1085)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou, Tsukuba, Ibaraki, Japan
E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170
Clones are derived from the mouse BAC library MSMG01. For BAC library availability, please contact Kunya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koydai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TU
LIBRARY : PBAC3.6
Vector : EcoRI
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1.1085
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-170B12.TU"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

FEATURES
source

1.1085
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-170B12.TU"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"

ORIGIN

Query Match 24.1%; Score 37.8; DB 9; Length 1085;
Best Local Similarity 51.9%; Pred. No. 0.46; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 CTACATGCTTGTCAACCGTCTGCTGTTGCATGCTCAAGCCTTGCTTATCGAAC 81
DB 719 CTCCT 660
QY 82 AAGAGATACCTACTCCCAACATTCATCTTACTGCAACTTCATGCAACACGCA 141
DB 659 ACATNATATACNCCCCCNANNAACANNAACCCGCAATTTTANANNAATAAA 600

QY 142 CATATGTTCCG 154
DB 599 CAAACTTTTGTG 587

RESULT 7
BB505306

LOCUS BB505306 700 bp mRNA linear EST 25-OCT-2001
DEFINITION BB505306 RIKEN full-length enriched, 10 days lactation, adult female mammary gland Mus musculus cDNA clone D730002G06, mRNA sequence.
ACCESSION BB505306
VERSION BB505306.2 GI:16442791
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
On Jul 27, 2000 this sequence version replaced gi:5914268.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamataka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES

source

1..700
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D730002G06"
/sex="female"
/tissue_type="mammary gland"
/dev_stage="10 days lactation, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days lactation, adult female mammary gland"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia

ORIGIN

Query Match 21.7%; Score 34; DB 2; Length 700;
Best Local Similarity 54.9%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 27 TCGTTGTTCAACCGTTGCTGTTGTCATCGCCAAAGCCTTGCTTGTGAACCAAG 86
DB 195 TTTCTTTTTCCTCCCTCCCTCTGATTTCTTGACCTCACTTTCATTTGGCTTTAAGT 254
QY 87 GATACCTACTCCCAACAATCCATCTTACTGCAACTTCATGCAACGACATAT 146
DB 255 GTTCTAGATTCCTCAAGAAAGCATTGTGACATGACGTCAACATATCAACCTGAAAT 314
QY 147 GT 148
DB 315 GT 316

RESULT 8
CL026582 1677 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-23G24.596.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-23G24, genomic survey sequence.
ACCESSION CL026582
VERSION CL026582.1 GI:40470443
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopus; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1677)
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
A physical map of the xenopus tropicalis genome
Unpublished (2003)
CONTACT: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: SP6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 125
High quality sequence stop: 217.
Location/Qualifiers

FEATURES

source

1..1677
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-23G24"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match 21.4%; Score 33.6; DB 9; Length 1677;
 Best Local Similarity 53.9%; Pred. No. 12;
 Matches 69; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 3 TGTGATCTCATGAGCATCTACTACATGCTTGTTCGAACGGTTCGTTGTTCCATGTCACCA 62
 DB 769 TGTGATCTCATGAGCATCTACTACATGCTTGTTCGAACGGTTCGTTGTTCCATGTCACCA 828
 QY 63 GCGTTCGCTTATTTGGAACCAAGAGATCTACTCCCAACATTCATCTTACTCATGCA 122
 DB 829 TCATCACTATCTCTTACAGTTACATGTCACCTCTCTCACTTCTATATATTATTCAGACA 888
 QY 123 ACTTCAT 130
 DB 889 TCTTTCAT 896

RESULT 9
 A1744861 440 bp mRNA linear EST 21-JUN-1999
 LOCUS cr16506.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218451 3',
 DEFINITION mRNA sequence.
 ACCESSION A1744861
 VERSION A1744861
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 440)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 www-bio.liml.gov/db/ftp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 411.
 Location/Qualifiers
 1..440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2218451"
 /issue_type="tumor, 5 pooled (see description)"
 /lab_host="DH10B"
 /clone_1lb="NCI_CGAP_Ov23"
 /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.35 kb. Tumor types include: mixed
 Mullerian tumor, papillary serous, clear cell, spindle
 cell. All are primary tumors, metastasis positive. Life
 Technologies catalog #: 11534-013"

ORIGIN
 Query Match 20.9%; Score 32.8; DB 1; Length 440;
 Best Local Similarity 56.5%; Pred. No. 14;
 Matches 61; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 46 CTTGTTTCATGTCGAACCTTCCCTATTTGAACCAAGATCTTACTCCCAACAA 105
 DB 259 CTTTGCCAGCAACCAATGATATCAATTCCTCCCAAGCAACACTCTTACTTATTT 318
 QY 106 TCCATCTTACTCATGCAACTTCATGCAACGACCATATGTTTCT 153

DB 319 TCCATCTTAAATTAACAGCGTCGATGTCCTAGATATCTGTTCT 366

RESULT 10
 CR575123 804 bp mRNA linear EST 21-JUL-2004
 LOCUS CR575123 XGC-tailbud-head Xenopus tropicalis cDNA clone THDA024m23
 DEFINITION 5', mRNA sequence.
 ACCESSION CR575123
 VERSION CR575123
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 1 (bases 1 to 804)
 Crouching, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (2004)
 JOURNAL Unpublished (2004)
 COMMENT Contact: Crouching MDR
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE_ID: THDA024m23.p1k5p6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Nigel Garrett.
 Seq primer: Sp6.
 Location/Qualifiers
 1..804
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="THDA024m23"
 /dev_stage="tailbud-head (stage 28-30)"
 /lab_host="Escherichia coli DH10B."
 /clone_1lb="XGC-tailbud-head"
 /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
 was oligo dT primed from 5ug of poly A+ RNA from tailbud
 head. EcoRI-NotI cut cDNA was then ligated into pCS107
 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
 Query Match 20.8%; Score 32.6; DB 7; Length 804;
 Best Local Similarity 63.3%; Pred. No. 20;
 Matches 50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 74 TCTGAACCAAGAGATCTTACTCCCAACATTCATCTTACTCATGCACTTCCATGCA 133
 DB 480 TCTGAATTTGAAAGGCCATCTCCCATGATTCATTTATGCAAGTATTTACATTTT 539
 QY 134 AACAGCATATGTTTCC 152
 DB 540 AAAAGTATTTTTCCTCC 558

RESULT 11
 CNS003FV/c 1101 bp DNA linear GSS 03-JUN-1999
 LOCUS CNS003FV/c Drosophila melanogaster genome survey sequence ITB3 end of BAC #
 DEFINITION BACR08003 of RPK1-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL064657
 VERSION AL064657
 KEYWORDS GSS.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1101)
 Genoscope.


```
FEATURES
source
location/Qualifiers
1..347
/organism="Mus musculus"
/mol_type="mRNA"
/id_xref="taxon:10090"
/clone="IMAGE:346691"
/sex="female (lactating)"
```

ORIGIN	Query Match	20.6%	Score 32.4	DB 1	Length 483
	Best Local Similarity	55.3%	Pred. No. 20		
	Matches 63	Conservative 0	Mismatches 51	Indels 0	Gaps 0
QY	21	ACTACATGCTTGTTCACACGCTTCGCTTGTTCATCGTCCAGCCTTGCTATTGTGAAC	80		
b6	371	ACACACATCTATATGATTAAGCTTTTATTTTAAAAAAGACCCCAATATATCTTAAC	312		

QY 81 CAGAGACTACTCTCCCAACATTCCTACTCATGCACTTCGATGCA 134
DB 311 ACAGTAAAGTGCACCTATCATCTAGTTAAATGAATATTTCCAGGAA 258

RESULT 14
LOCUS AW990724/c 497 bp mRNA linear EST 02-JUN-2000
DEFINITION ufl10b11.x1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
ACCESSION AW990724
VERSION AW990724.1 GI:8186451
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: ufl10b11.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-rc@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:937793
High quality sequence stop: 459.
Location/Qualifiers
1..497
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1510941"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 497;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TCGTTGTTCAACCGTTCGTTGTTGTCATGTCGCAAGCTTGCTATTGCAACCAAG 86
DB 490 TTTCTTTTTCACCTCCCTCTGATTTCTTTGTACTTCATCTTATTTGGCCCTTTAGTT 431

QY 87 GATACCTACTCCCAACATTCATCTTACTCATGCACTTCGATGCAACAGCAGACATAT 146
DB 430 GTTCTAGATTCCCAAGAGCATTGTCACCTAGCTCAAGTCAACATTAACACCTGAAAT 371

QY 147 GT 148
DB 370 GT 369

RESULT 15
LOCUS BE632212/c 498 bp mRNA linear EST 25-AUG-2000
DEFINITION uul1f01.x1 Soares mammary_gland_NMLMG Mus musculus cDNA clone
ACCESSION BE632212
IMAGE:3371641 3', mRNA sequence.

VERSION BE632212.1 GI:9914830
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: uul1f01.y2
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-rc@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1081245
High quality sequence stop: 459.
Location/Qualifiers
1..498
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3371641"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 20.6%; Score 32.4; DB 2; Length 498;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TCGTTGTTCAACCGTTCGTTGTTGTCATGTCGCAAGCTTGCTATTGCAACCAAG 86
DB 491 TTTCTTTTTCACCTCCCTCTGATTTCTTTGTACTTCATCTTATTTGGCCCTTTAGTT 432

QY 87 GATACCTACTCCCAACATTCATCTTACTCATGCACTTCGATGCAACAGCAGACATAT 146
DB 431 GTTCTAGATTCCCAAGAGCATTGTCACCTAGCTCAAGTCAACATTAACACCTGAAAT 372

QY 147 GT 148
DB 371 GT 370

Search completed: September 16, 2005, 08:08:43
Job time : 490.661 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 100.667 Seconds
(without alignments)
14440.280 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaaccgttcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_srb:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	30	100.0	30	6	AX224397 Sequence
2	30	100.0	255	6	AX224402 Sequence
3	30	100.0	1394	6	AX224394 Sequence
4	30	100.0	1394	6	AX224395 Sequence
5	30	100.0	1394	6	BD062176 Male t1ss
6	30	100.0	1394	6	BD062177 Male t1ss
7	30	100.0	3343	6	AF360356 Zea mays
8	24.2	80.7	158	6	AX224396 Sequence
9	22	73.3	1304	6	AX523502 Sequence
10	21.6	72.0	187192	2	AC124638 Mus muscu
11	21.6	72.0	188427	2	AC151282 Mus muscu
12	21.2	70.7	129778	8	AC091123 Oryza sat
13	21.2	70.7	140729	8	AP006548 Oryza sat
14	21.2	70.7	153675	2	AP005710 Oryza sat
15	21.2	70.7	155146	8	AP004118 Oryza sat
16	21.2	70.7	158374	8	AP004786 Oryza sat
17	21.2	70.7	172058	8	AP003682 Oryza sat
18	21.2	70.7	178158	8	AC084320 Oryza sat
19	21	70.0	91298	2	AC119671 Oryza sat

C	20	21	70.0	93342	9	AC087245 Homo sapi
C	21	21	70.0	123472	8	AC120984 Oryza sat
C	22	21	70.0	142015	8	AC120884 Oryza sat
C	23	21	70.0	143681	8	AC082645 Oryza sat
C	24	21	70.0	156394	2	AC026558 Homo sapi
C	25	21	70.0	158839	8	AC090882 Oryza sat
C	26	21	70.0	162198	8	AP005916 Oryza sat
C	27	21	70.0	177790	8	AP005930 Oryza sat
C	28	21	70.0	187707	8	AC077693 Oryza sat
C	29	21	70.0	188854	10	AL626770 Mouse DNA
C	30	21	70.0	253492	2	AC102160 Mus muscu
C	31	21	70.0	300029	8	AE017121 Oryza sat
C	32	20.6	68.7	173088	8	AC109365 Oryza sat
C	33	20.6	68.7	196303	10	AC101349 Mus muscu
C	34	20.6	68.7	206372	10	EX005189 Mouse DNA
C	35	20.4	68.0	1026	14	H1M429890 Human 1mm
C	36	20.4	68.0	1542	6	AX506929 Sequence
C	37	20.4	68.0	1762	8	AY049275 Arabidops
C	38	20.4	68.0	1682	8	AT925213 Arabidops
C	39	20.4	68.0	1913	8	AF242307 Euphorbia
C	40	20.4	68.0	1985	8	AC067922 Homo sapi
C	41	20.4	68.0	42868	9	AC115980 Mus muscu
C	42	20.4	68.0	65406	2	AC021665 Arabidops
C	43	20.4	68.0	80413	8	OSJN00036 Arabidops
C	44	20.4	68.0	114051	8	AL606598 Oryza sat
C	45	20.4	68.0	146776	8	AP007223 Oryza sat

ALIGNMENTS

RESULT 1	AX224397	Sequence 4 from Patent WO0160997.	30 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224397					
DEFINITION	AX224397					
ACCESSION	AX224397					
VERSION	AX224397.1	GI:15554639				
KEYWORDS						
SOURCE						
ORGANISM						
Zeae mays						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.						
REFERENCE						
AUTHORS	Albertsen,M.C., Fox,T.W., Garmaat,C.W., Huffman,G. and Kendall,T.L.					
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 4 23-AUG-2001;					
PIONEER	HI-BRED INTERNATIONAL, INC. (US)					
FEATURES						
source		1..30				
		/organism="Zea mays"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:4577"				
ORIGIN						
Query Match		100.0%;	Score 30;	DB 6;	Length 30;	
Best Local Similarity		100.0%;	Pred. No. 0.03;			
Matches	30;	Conservative	0;	Mismatches	0;	Indels 0;
				Gaps	0;	
Qy	1	CATGCTGTTCAACCGTTCGCTTGTTC	30			
	1					
Db	1	CATGCTGTTCAACCGTTCGCTTGTTC	30			
	1					
RESULT 2	AX224402	Sequence 9 from Patent WO0160997.	255 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	AX224402					
ACCESSION	AX224402					
VERSION	AX224402.1	GI:15554644				
KEYWORDS						
SOURCE						
ORGANISM						
Zeae mays						

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68
|||||

RESULT 3
AX224394 1394 bp DNA linear PAT 10-SRP-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0160997.
AX224394
VERSION
AX224394.1 GI:15554636
KEYWORDS
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
|||||

RESULT 4
AX224395 1394 bp DNA linear PAT 10-SRP-2001
LOCUS
DEFINITION
Sequence 2 from Patent WO0160997.
AX224395
VERSION
AX224395.1 GI:15554637
KEYWORDS
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1 Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
|||||

RESULT 5
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
Male tissue-preferred regulatory region and method of using same.
BD062176
VERSION
BD062176.1 GI:22607781
KEYWORDS
JP 2001520523-A/1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAA,T,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/62,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K4/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH key Location/Qualifiers

FEATURES
source
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
|||||

Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
|||||

RESULT 6
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
Male tissue-preferred regulatory region and method of using same.
BD062177
VERSION
BD062177.1 GI:22607782
KEYWORDS
JP 2001520523-A/2.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI-BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
RESULT 7
LOCUS AF360356 3343 bp DNA linear PLN 12-MAY-2001
DEFINITION Zea mays male fertility protein (Me45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
TITLE Cloning of Me45, a gene required for male fertility from Zea mays
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3343)
AUTHORS Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
FEATURES
source Location/Qualifiers
1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivars="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Me45"
/contig="Me45"
/join(1392..1768,1898..2182,2280..2447,2534..>2942)
/product="male fertility protein"
/join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Me45"
/codon_start=1
/product="male fertility protein"
/protein_id="AK52489.1"

/db_xref="GI:14028757"
/translation="MEKRLQRRRGDGIYQYBHLFFALALALVADPGSLAEV
DYRPVHLELAPGEVWGSWPRDNASKRLRGRLEFGEVGPESIEFDLQGRPYGLA
DQVVMGEBAEWETFAVMNPDWSEVCANVSTTRQHEKEBFCGRPLGRFGE
TSELVYADAYVIGLVVQSGVAA SVARADSDP IRFANDLVHNGSVFTDSMKY
SKRDHLNILEEGGTGRLRYDPETSGVAVVKGIVFPNGVOISDHQPLFSETTNC
RIMRWVLGEPRAGEVFPANLPGFPDNPVNSNGCGPWWAIDCCRTPAQGVFAKRPRLR
TIYKRPFLSKVLVIMKARRMHTVLLALDDEGRVVEVLEDRGHVVKLVSEVREGRK
LWIGTVAHNHVIATIPYLED"

ORIGIN
Query Match 100.0%; Score 30; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
RESULT 8
LOCUS AX224396 158 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source Location/Qualifiers
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 80.7%; Score 24.2; DB 6; Length 158;
Best Local Similarity 89.7%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTTGTTCAACCGTTCGTTGTTCC 29
Db 25 CATGCTTGTTCAACCGTTCGTTGTTCC 53
RESULT 9
LOCUS AX523502 1304 bp DNA linear PAT 24-OCT-2002
DEFINITION Sequence 90 from Patent WO02064788.
ACCESSION AX523502
VERSION AX523502.1 GI:24412398
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1
AUTHORS Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
TITLE Compositions and methods relating to lung specific genes and
proteins
JOURNAL Patent: WO 02064788-A 90 22-AUG-2002;
Diadexus, Inc. (US)
FEATURES
source Location/Qualifiers
1..1304
/organism="Homo sapiens"


```

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

```

Query Match	73.3%	Score 22	DB 6	Length 1304
Best Local Similarity	83.3%	Pred No. 97		
Matches 25, Conservative	0	Mismatches 5	Indels 0	Gaps 0

```
QY      1 CATGCTTGTCACCGTTCGTCTGTGCCA 30
          |||||
DB      1155 CTTCTGTGTCCTCCCGTCTCTGTGCCA 1126
```

RESULT 10	AC124638/c	LOCUS	DEFINITION
AC124638	187192 bp	DNA	linear
Mus musculus chromosome 17 clone RP23-461H1 map 17, *** SEQUENCING			
IN PROGRESS ***, 10 unordered pieces.			

ACCESSION AC124638
VERSION AC124638.3 GI:45120281
KEYWORDS HTG, HTGS PHASE1, HTGS_PULLTOP, HTGS_ACTIVEFIN
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 187192)	Birren, B., Nussbaum, C. and Lander, E.	Mus musculus chromosome 17, clone RP23-461H1

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 187192)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,

Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Ioh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McDonald, D.

Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, D., Roman, I., Rosetti, M., Roy, A., Sattar, D., Schwan, C.

Stojanovic, N., Strause, N., Sudriamman, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, I., Zimmer, A. and Zody M.

BOOKEND
Submitted (15-JUN-2024), Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 187192)
AUTHORS
Birren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N.,

Collimore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,

Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,

O'Neil, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Rachupya, A., Ramasaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.

Talanas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.J., Zembek, L., Zimmer, A. and Zoody, M.

Submitted (05-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 5, 2004 this sequence version replaced gi:28893670.

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information

Center project name: L2411
Center clone name: 461_H_1

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
*
* -----

```

* as soon as it is available and the accession number will
* be preserved.
*      1      9926: contig of 9926 bp in length

```

9927	10026: gap of 100 bp
10027	20081: contig of 10055 bp in length
20082	20181: gap of 100 bp

*	43360	45947: contig of 2588 bp in length
*	45948	46047: gap of 100 bp
*	46048	100940: contig of 54893 bp in length

*	106720	106819:	gap of 100 bp
*	106820	140370:	contig of 33551 bp in length
*	140371	140470:	gap of 100 bp

Accession	Contig	Length (bp)
147806	contig of 4760 bp in length	4760
152565	contig of 4760 bp in length	4760
152566	gap of 100 bp	100
152665	contig of 34527 bp in length	34527
152666	contig of 34527 bp in length	34527

```
/organism="Mus musculus"  
/mol_type="genomic DNA"  
/db_xref="taxon:10090"  
/chromosome="17"
```

ORIGIN
/clone="RP23-461H1" /clone_lib="RPCT-23 Female Mouse BAC"

by	2	ATGCTTGTTCACCGTTCGCTCTGTTCC	29				
		Matches	24;	Conservative	0;	Mismatches	4;
						Indels	0;
						Gaps	0

RESULT 11

DEFINITION	Mus musculus chromosome 17 clone Rn23-461H1, *** SEQUENCING IN
PROGRESS	***, 12 unordered pieces.
ACCESSION	AC151282

VERSION	AC151282.1	GI:51511039
---------	------------	-------------


```

KEYWORDS      HTG: HTGS_PHASE1.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         1 (bases 1 to 188427)
JOURNAL       Wilson, R.K.
REFERENCE     The sequence of Mus musculus clone
AUTHORS       2 (bases 1 to 188427)
TITLE         Unpublished
JOURNAL       Wilson, R.K.
REFERENCE     Direct Submission
AUTHORS       Submitted (24-AUG-2004) Genome Sequencing Center, 4444 Forest Park
TITLE         Parkway, St. Louis, MO 63108, USA
JOURNAL
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
----- Project Information -----
Center project name: M_BA0461H01
----- Summary Statistics -----
Sequencing vector: MJ3, 0%
Sequencing vector: plasmid, 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 18266 bases at least Q40
Consensus quality: 183763 bases at least Q30
Consensus quality: 184341 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1129: contig of 1129 bp in length
* 1130 1229: gap of unknown length
* 1230 2832: contig of 1603 bp in length
* 2833 2932: gap of unknown length
* 2933 4658: contig of 1726 bp in length
* 4659 4758: gap of unknown length
* 4759 6427: contig of 1669 bp in length
* 6428 6527: gap of unknown length
* 6528 11124: contig of 4597 bp in length
* 11125 11224: gap of unknown length
* 11225 21177: contig of 9953 bp in length
* 21178 21277: gap of unknown length
* 21278 31319: contig of 10042 bp in length
* 31320 31419: gap of unknown length
* 31420 54758: contig of 23339 bp in length
* 54759 54858: gap of unknown length
* 54859 95804: contig of 40946 bp in length
* 95805 95904: gap of unknown length
* 95905 135186: contig of 39282 bp in length
* 135187 135286: gap of unknown length
* 135287 186683: contig of 51397 bp in length
* 186684 186783: gap of unknown length
* 186784 188427: contig of 1644 bp in length.
*
* Location/Qualifiers
* 1.188427
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /chromosome="17"
* /clone="RP23-461H1"
* 1.1129
* /note="assembly_Contig11"
* 1230. .2832

```

Query Match	Best Local Similarity	85.7%	Pred. No. 1.4e+02	Matches 24	Conservative 0	Mismatches 4	Indels 0	Gaps 0
Qy	2	ATGCTGTTCACCGTTCGCTGTGTTCC	29					
Db	182832	ATGCTGTTCACAGTTCATCTGTTC	182859					
RESULT 12	AC091123/c	LOCUS	AC091123	129778	bp	DNA	linear	PLN 09-AUG-2001
DEFINITION	Oryza sativa chromosome 3 BAC OSJNB0093E13	complete sequence.						
ACCESSION	AC091123	VERSION	AC091123.4	GI:14670090				
KEYWORDS	HTG.							
SOURCE	Oryza sativa (japonica cultivar-group)							
ORGANISM	Oryza sativa (japonica cultivar-group)							
REFERENCE	Euharvota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.							
AUTHORS	1 (bases 1 to 129778)							
REFERENCE	Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N., Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M., Taitzin,T., Riggs,F., Hsiao,J., Ziemann,V., Blunt,S., Pal,G., Vavilov,S.E., Utecherack,T.R., Feldlyum,T.V., Quackenbush,J., Salzberg,S.L., White,O. and Frazer,C.M.							
TITLE	Oryza sativa chromosome 3 BAC OSJNB0093E13 genomic sequence							
REFERENCE	Unpublished							
JOURNAL	2 (bases 1 to 129778)							
TITLE	Buell,R.							
AUTHORS	Direct Submission							
REFERENCE	Submitted (31-MAR-2001) The Institute for Genomic Research, 9712							
JOURNAL	Medical Center Dr, Rockville, MD 20850, USA							
TITLE	3 (bases 1 to 129778)							
AUTHORS	Buell,R.							
REFERENCE	Direct Submission							
JOURNAL	Submitted (11-JUL-2001) The Institute for Genomic Research, 9712							
TITLE	Medical Center Dr, Rockville, MD 20850, USA							
REFERENCE	4 (bases 1 to 129778)							
AUTHORS	Buell,R.							
REFERENCE	Direct Submission							
JOURNAL	Submitted (14-JUL-2001) The Institute for Genomic Research, 9712							
TITLE	Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org							
REFERENCE	5 (bases 1 to 129778)							
AUTHORS	Buell,R.							
REFERENCE	Direct Submission							
JOURNAL	Submitted (18-JUL-2001) The Institute for Genomic Research, 9712							
TITLE	Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org							
REFERENCE	Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org							

REFERENCE 6 (bases 1 to 129778)
AUTHORS Buell, R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Jul 11, 2001 this sequence version replaced gi:1404299.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0093E13 is from *Oryza sativa* chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC
clone.
Genes were identified by a combination of several methods: Gene
prediction programs including fgenesh (<http://www.softberry.com/>),
genscan and Genscan+ (Chris Burge,
<http://CCR-081.mit.edu/GENSCAN.html>), GeneMark-ES (Mark Borodovsky,
<http://genemark.biology.gatech.edu/GeneMark/>), and Gensplicer
(Mihela Perrea and Steven Salzberg, contact.mperrea@tigr.org),
searches of the complete sequence against a peptide database and
the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.shtml>).
Annotated genes are named to indicate the level of evidence for
their annotation. Genes with similarity to other proteins are named
after the database hits. Genes without significant peptide
similarity but with EST similarity are named as unknown proteins.
Genes without protein or EST similarity, that are predicted by more
than two gene prediction programs over most of their length are
annotated as hypothetical proteins. Genes encoding tRNAs are
predicted by tRNAscan-SE (Sean Eddy,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are
identified by RepeatMasker (Arian Smit,
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>).
This BAC overlaps with rice BACs OSJNBa0091J19 (GB:AC084320) and
OSJNB0024U04.

FEATURES
source 1..129778
Location/Qualifiers
1..129778
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/map="near S13014"
/clone="OSJNB0093E13"
1701..1821
/rpt_family="Gaijin Oa2 MITE element from gb:D10985 Rice
seirine cbp3 gene for carboxypeptidase III (1125 to 1251)
127 nt"
complement(2452..2476)
repeat_region
/rpt_family="(GAA)n"
4620..5435
/gene="OSJNB0093E13.4"
/note="predicted by fgenesh"
join(<4620..4718,4866..5435)
/gene="OSJNB0093E13.4"
join(4620..4718,4866..5435)
/gene="OSJNB0093E13.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72893.1"
/db_xref="GI:14718315"
/translation="MEDWVLSDSQDSVLEHDSGVGSGDSGFAARSGKLPSSSO
LILTLKSTQPKRTRRTLPVSPALAGGSHENAEISDVARNNDHVSNIAAEDTTF
SGEEDLDDTDEIDECFDEEDGICENPDDEIDPDDEBELDDDESDIDDICEDYEDK
ICEENPDDEIDFDEEEDICEEDLDDDDCESLDDNDICFPADIEDIICLETRTKSST
NIAM"
4640..4775
repeat_region
/rpt_family="(CGG)n"
complement(4730..4835)
repeat_region
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
411008..511724
mRNA
/gene="OSJNB0093E13.5"

CDS
11008..11724
/gene="OSJNB0093E13.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72895.1"
/db_xref="GI:14718317"
/translation="MEVRRRAVAESGRRRQAYGEOROSGRROLATGRLALPEA
KGRWSEGERKATVSGRALGVRGASASSASAGRGGRARVYRHTPTAAPVSV
FGDGRROEMGGNGVYBARSKTILDMGRPLSIGFIQWGRRCGGDIDGRLQGW
RAGAYKMAAGAHGRORSSSGGKKKGRVRRGSGTPTMGREGREMLMVLVLDASA
QKMGAGHGHRGAGCGQ"
12209..12383
repeat_region
/rpt_family="(TACA)n"
complement(12766..16846)
/gene="OSJNB0093E13.7"
/note="highly similar to retrofit gag pol protein
gb:AB82754 GI:2586082 (Oryza longistyliseta)"
complement(join(<12766..12903,13103..13518,13741..16846))
/gene="OSJNB0093E13.7"
complement(join(12766..12903,13103..13518,13741..16846))
/codon_start=1
/product="putative gag-pol protein"
/protein_id="AAK72882.1"
/db_xref="GI:14718304"
/translation="MASSSSSTAFLNPLFGVQVSEKLTQKNSLSAQLTALRGARL
EGHVTGVPPEALEQKEGKEKTVAVPNPAYGEWTFATDQVGLPSSLTREIS
OYAGAPTPAAAKTKTEMTFSRRSGAINVRLATTTQCGOSTVTEYESSKMALGDEI
NATGCTPDEELVATITNGLSRFPVYBALAKRASAIVAAVYSQILGFENPKVIRT
ACAAVTSRGGSGNQGRGNGPRGTGSGGGRGGHGRNGRGGTNNKPTQVC
HKKGTVADCMWRVDENYVPDEKLGAAVTHAYGVDTNNYVDTEADH1TGQJDLKTR
EYKGTQD1HTASGEAMD1OH1GHSVYPTSRPLK1LHVPKASK1L1SVRLVAD
NYAFLEIHKQKFL1KDVTRRTILEGPCRBD1YLPADPTIKQVAVMSPFRMGR1L
GASKEP1LV1NOMK1LPCSNPSESGSCDCCGCKSHOLPEPKFSSVNLTE1IHS
DVMGPASVGAARYVFPIDYKFTWIT1YLFKSBVPEKKEQSNVEROPNKL1L
GMQTDWGEYQ1LNSFFKQ1G1AHQVSCPHHQOQGAVERKRI1VE1GLSLA1AS
PLK1WDEAF1LATY1L1N1LPRKVI1DFDPLRL1FQOPDYSLRTFGCA1CWN1LAPY
THK1QFRSKQCV1LGF5N1LHKFKCLEVATQCV1YSRVTPEQCFPF5K1QPN1AR
1LRAE1SL1PAD1TSHASDQGVAAHNDHMH1L1PVYTSFPENDAADT1DADSNAE1
NAEOPDGIHPAS1GCVHTA1ETREVA1ESPRTDPA1VOA1PAPVTPA1AAS1GAV
SSRASRS1ASHQ1GDTSP1P1PHHD1AGSSA1EPAV1ATEBGD1NNA1PKTR1QSG1R
KK1YTDGT1VRYGCTSS1GEPON1LAEL1AGDRNWK1EAK1EATK1NKT1WH1VPRK1G
IN1IDCKWYK1K1RAKADSLR1YKRL1AKAKQK1GYG1DEBTF1SPVKA1AT1RVL1S
LVAS1KDTSL1PFYDKK1GVFM1LVVD1L1VSS1SSKAS1AL1QD1NOF1AL1DL1DL1H
FPL1GEYK1NDGL1V1TQEKY1AD1L1RVMS1D1G1CPVNT1YS1SK1L1SOGGAL1GPN
D1STOYRS1V1GAL1OY1TL1R1P1S1SVNN1DPT1LY1G1VHNK1Q1FLDQV1L1K1NTR1HM1Q
QOK1CY1K1YFRN1LV"
complement(13497..13708)
repeat_region
/rpt_family="Rice retrotransposon Tos7 DNA for reverse
transcriptase, partial cds gi|1621482|dbj|D85871.1|D85871"
complement(16049..16146)
repeat_region
/rpt_family="(CGG)n"
17933..19954
/gene="OSJNB0093E13.6"
/note="predicted by fgenesh"
417933..519954
mRNA
/gene="OSJNB0093E13.6"
17933..19954
CDS
/gene="OSJNB0093E13.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72884.1"
/db_xref="GI:14718306"
/translation="MEDWVLSGSGGDSVLEHDSGVGSGDSGFAARSGKLPSSSO
LILTLKSTQPKRTRRTLPVSPALAGGSHENAEISDVARNNDHVSNIAAEDTTF
SGEEDLDDTDEIDECFDEEDGICENPDDEIDPDDEBELDDDESDIDDICEDYEDK
ICEENPDDEIDFDEEEDICEEDLDDDDCESLDDNDICFPADIEDIICLETRTKSST
NIAM"
4640..4775
repeat_region
/rpt_family="(CGG)n"
complement(4730..4835)
repeat_region
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
411008..511724
mRNA
/gene="OSJNB0093E13.5"

AM00AARNOAAAMVRAADADSDYVQAAAGSLA0DVLEALSGCATGLRMAAPSPA
VGVDVASNTAAAFLEPDTALADGIVSLAVFYLLRGVYLLRLRICALN"
repeat_region
18749..18856
/rpt_family="(GGA)n"
repeat_region
complement(20115..20185)
/rpt_family="AT_rich"
gene
21519..21947
/gene="OSJNB00093E13.8"
/note="predicted by Egeneeh"
mRNA
<21519..>21947
/gene="OSJNB00093E13.8"
21519..21947
/gene="OSJNB00093E13.8"
/codon_start=1
/product="hypothetical protein"
Query Match 70.7%; Score 21.2; DB 8; Length 129778;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 5 CTGTTCACACCGTTCCTGTTCCTCA 30
Db 20209 CATGTTTACCGTTCCTGTTCCTCA 20184
RESULT 13
AP006548 140729 bp DNA linear PLN 12-JUL-2003
LOCUS
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9,
DEFINITION
AP006548
PAC clone: P0705E11.
ACCESSION
AP006548.1 GI:32490470
KEYWORDS
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
SOURCE
ORGANISM
Oliveira,A.C., Matos,L.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,
Malone,G., and Dellagostin,O.
Oryza sativa japonbare genomic DNA, chromosome 9, PAC
clone:P0705E11
Published Only in Database (2003)
2 (bases 1 to 140729)
Oliveira,A.C., Matos,L.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,
Malone,G., and Dellagostin,O.
Direct Submission
Submitted (08-JUL-2003) Antonio C Oliveira, Universidade Federal de
Pelotas, Centro de Genomica e Fitomelhoramento, Department
Fitotecnica; Campus Universitario do Capao do Leao, Pelotas,
Pelotas, Rio Grande do Sul 96010900, Brasil
(E-mail:acostol@terra.com.br,
URL:http://www.ufpel.br/faem/fitotecnica/fitomelhoramento/
Tel:55-53-275-7263, Fax:55-53-275-9031)
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), Glimmer
(http://www.tigr.org/tdb/glimmer/glmr.form.html), RiceHM
(http://xsp.dmr.affrc.go.jp/RiceHM/), SplicePredictor
(http://bioinformatics.lestate.edu/cgi-bin/bp.cgi), BLASTN and
BLASTX. The genomic sequence was searched against NCBI NonRedundant
Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the EST sequence database at
RGP. Protein homologues of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP. ESTs represent the
identified cDNA sequences using BLASTN with the corresponding DBJ
accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering

almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted by two or more gene prediction
programs is classified as a 'hypothetical' protein according to
IRGSP standard.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
This sequence of P0705E11 clone has an overlap with OSJNB0069D16
clone (DBJ: AP005882) at 5' end and an overlap with OSJNB0069D16
clone (DBJ: AP005862) at 3' end.
FEATURES
source
1..140729
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="japonbare"
/db_xref="taxon:39947"
/chromosome="9"
/clone="P0705E11"
complement(1590..2675)
/note="3'LTR"
3382..4071
/gene="P0705E11.1"
3382..4071
/gene="P0705E11.1"
/note="probably inactive due to 3' exon missing in CDS
probably inactive due to 5' exon missing in CDS
pseudogene, orf3 of gypsy-type retrotransposon RIR2"
/pseudo
complement(7148..9310)
/gene="P0705E11.2"
complement(7148..9310)
/gene="P0705E11.2"
/note="probably inactive due to 5' exon missing in CDS
pseudogene, GAG-POL precursor of gypsy-type
retrotransposon RIR2"
/pseudo
complement(13300..14385)
/note="5'LTR"
20095..21408
/gene="P0705E11.3"
20095..21408
/gene="P0705E11.3"
/note="start and end point are not identified"
20095..21408
/gene="P0705E11.3"
/codon_start=1
/product="putative anthranilate N-benzoyltransferase"
/protein_id="BAC79154.1"
/db_xref="GI:32490471"
/translation="MEVYVLSKIVKPAVNGVAAAPVVEYIPLSIFPKVYKQMAI
IYAFPPAPSTAIETKGLAVLAOYRAFRGQSGSPDGEAAVVLNDRKRLVEAAVDA
DIVDAAPAKPTPELRLHPDLGEQEVVLLQLTFRGCSLAVGFTSNHNVADGHAIS
NFLVAMGATRCGLPMGAPVNHIALPKRSPVHEHNRNEVYLPAAGDSHGHD
GGAANIVYHKAHFTKDFIAGLRAAASGGRSPFRFETILAHLMRTMTRAGLSHDE
ASTIRLSVDRRLGAPAEVGNVLMAPRAVTDGLLTRPLKHAAYVHDEVARVDG
AYFRFLRPAISGAGDGEGLRASVILVDVCPNMEVSDSLTFPPYEYLDFTGSGPTTF
MESTPTGMLFLVPSYIYGDGSDVDAFVFFVFNHLEAFKCYTSM"
33065..34423
/gene="P0705E11.4"
33065..34423
/gene="P0705E11.4"
/note="start and end point are not identified"
33065..34423
/gene="P0705E11.4"
/codon_start=1
/product="putative anthranilate N-benzoyltransferase"
/protein_id="BAC79155.1"
/db_xref="GI:32490472"
/translation="MEVYVLSRLVRSVPASAAPEEPSSMFDKTYTDMQMAIT
IAFRPPGSVADIERKGLAVGVPLFAGQVVRGGGLRGLVNDHNDHVAEACVDG
SLADIAPAKPSPVVLRHPSLGEIEVVOQLTFACGSLAVGTANHAAVADGATS
DFLVAMGAAAGLAVALATAAPRHHNPMFPRDPLVEFSGVRYVRRPPRAGV
GVGDGQHQHGGEBASHGIVIHAKFTDQFIRLRAASSEGGRFSPRETIAL
VWRMTMTRAGISNPLOSTTIRSVDRGLSAPAGVFNVLVMAFPRATVGDILGRPL
KHAQVYHDAVARADAAIFRSFVDPASSGAVGEGLAFTAVLVKYLCPDLEVDSDWITF

Gene	misc_feature
Gene	42989..45252
	/gene="P0705E11.5"
	42989..45252
	/gene="P0705E11.5"
	/note="Anticirrhinum majus shows structural homology to the maize transposon En/Spm and has no sequence specificity of insertion
	probably inactive due to 3' exon missing in CDS
	probably inactive due to 5' exon missing in CDS
	probably inactive due to including frameshift(s) in CDS
	pseudogene, transposable element Tami"
	/pseudo
Gene	49006..52703
	/gene="P0705E11.6"
	join(<49006..50427,50502..51258,51322..51598,51708..51881
	51978..52088,52173..52245,52340..52438,52518..525703)
	/gene="P0705E11.6"
	/note="start and end point are not identified"
	join(49006..50427,50502..51258,51322..51598,51708..51881,
	51978..52088,52173..52245,52340..52438,52518..52703)
	/gene="P0705E11.6"
	/codon_start=1
	/product="reverse transcriptase - like protein"
	/protein_id="BAC79156.1"
	/db_xref="GI:32490473"
	/translation="MADRDKQIILVDITAESSSQYMEEGNEEDPNQYLINEGVNRD
	ASGNOGCHYEREGNOEEBASAQSPPSYGKORACGAKKLEGGHITVEVEDRP
	SAPAAAKYVRHSGMYRVDRVNVSTVYKRRTRARGHESFVDPSEKEMLTMTMLE
	TLPTGEDVSKWTLKRAEPOFQFSGDLYQKYLKQPTFPDFPELKRHDWEVAY
	KTGEQKAGMEENKEMNAKKRYHHHLGSGYSVAMPKQEMASLIERGIEPTANMP
	ESKSKYVAHGGELTLNPADGSLVFGYIOEARRLTDVBAASSQGTPEPDRDRELTAL
	LQTPHPRTRKGVIPWKIGFKEDIHLYRSMNSKDDTAKLATVAYEIRKSAKXY
	EVARKVERMAHNRSHDPOPTIPRPAWSPGNSCASNGOVSQMDMORQDST
	CEVDITDRQPELHIEPKNKISIVASAMAPIDPDSGTYKRPAGYSKVEYELPG
	AYEDILTFPGDGETHRLDTCGAILLMRRRYITLLPRQASRAPSPALPSPDPA
	ESPAPAPSPPOAPASTPQDPAPVTPPRATPTTPPAPLPAPSKSAPPAPPLAHR
	AKKVAKAAKNNNDPGDQTEBDLDAVVASVKQPFRRPEKKIPLDPSRNFEGM
	SAKSEVALKLSYERTLLKKAASSGKAPQIQEPNOEIEPLVTVGVAPEIKSAKXY
	ELGKLYDALDGLSPTOMYKFOHLYMMSATGSEMIGARLRDTPFLOGDILMINR
	GYEGLYQDALDVCIMSCWIMLEIIOBARRRRVPTGFLDPKXVAMLDOYQETED
	LVHLKADHYKRTFLLPYNTSEHYVLLIDLEACTVYVNSMKSESTPRKVEILDR
	AMYRRHLVRGKMERLRARRKRFPCAKQKQSTNLCGYVVCYCHCADQITITRELD
	IMRNRLTTHKEFLAAVQELMGFLINEIILDPKSEFYDQNTJHRSLSAEVLAISTTS
	KS"
Gene	57646..64290
	/gene="P0705E11.7"
	join(<57646..57935,58078..58192,59352..59443,59596..59709
	59896..59950,60105..60215,60983..61036,61119..61294,
	61380..61487,62089..62162,62257..62432,62928..63021,
	63291..63311,63470..63523,64003..64137,64210..64290)
	/gene="P0705E11.7"
	/note="start and end point are not identified"
	join(57646..57935,58078..58192,59352..59443,59596..59709,
	59896..59950,60105..60215,60983..61036,61119..61294,
	61380..61487,62089..62162,62257..62432,62928..63021,
	63291..63311,63470..63523,64003..64137,64210..64290)
	/gene="P0705E11.7"
	/note="containing EST(s) : AU101437 [E31530] , AU101438 [E31530] "
	/codon_start=1
	/product="putative serine/threonine-protein kinase crr1"
	/protein_id="BAC79157.1"
	/db_xref="GI:32490474"
	/translation="MAVEEALAESCGSHAAASAAGRRGGGGGGGATSSSSASAGA
	AAARAKQOOQORHLAEVTEALRLRHDSGEPAERKSGFDELMHNHNRLLPAPYADV
	INVERADLTHRLLEAGADPAORAPAVRAVQYSPLIDNQNDDBASNTGEASRL
	LNROOSHPPRFGSSTNLALALAEASQGDHSDSDNVRPMEHTPFSTIDPK
	ILSTSLTLGRLGNTIGCAHAPSTNDQSLDVPVSGMDEBEDLIESRKIKGID
	ETGKSTTHSSSVENKMOIGENASADHVEYTPDQASMEITDGLKIFGKVASGSYG
	DIYRQYCSQDAIKVLKPERINADMQRKEQEVYIKRKVHKNVQVFIACQKSPNL
	CLVTEVSSVAGSYDYLHKGKGVFKLPALGVMDVSGMSYVLIQNNTIHEDLTGAIL
	MDENVTVAAGFVAVKAQSGVTAETGTRMAAPVIEIKRQYDHKADVAFSGSIIM
CDS	

gene	ELTGTGTAATGAAAGVVGKGLRPTPKNAHKLSTLQKCMQGEBAKRPDPFS ITLTGRIAAEVEDHDEHDKKEKILGLGFSALGRGH"			
mRNA	/gene="P0705511.8" Join(<65979. .66134,66228. .66279,67599. .75119,75210. .757533. .76241,76395. .76475,76728. .77036,78187. .78573,78721. .79524,79613. .79975,80194. .80424,80691. .81035,81251. .81469,81516. .81572,81735. .82079,82295. .82513,			
Query Match	70.7%;	Score 21.2;	DB 8;	Length 140729;
Best Local Similarity	88.5%;	Pred. No. 2.2e+02;		
Matches	23;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
Db	55011 CATGTTAACCGTTCGCTTATTCCA 55036			
5	CTTGTTCAACCGTTCGCTTGTTCCA 30			
RESULT 14				
LOCUS	AP005710			
DEFINITION	AP005710 153675 bp DNA linear HTG 06-SEP-2002			
ACCESSION	AP005710.1 GI:22775481			
VERSION	HTG; HTGS PHASE2.			
KEYWORDS	Oryza sativa (japonica cultivar-group)			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.			
REFERENCE	1			
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.			
TITLE	Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC clone: P0646B07			
JOURNAL	Published Only in Database (2002)			
REFERENCE	2 (bases 1 to 153675)			
AUTHORS	Sasaki, T., Matsumoto, T. and Katayose, Y.			
TITLE	Direct Submission			
JOURNAL	Submitted (05-SEP-2002) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan			
COMMENT	/e-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-58-7441, Fax:81-298-58-7468)			
	NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.			
	* NOTE: This is a 'working draft' sequence.			
	* This sequence will be replaced			
	* by the finished sequence as soon as it is available and			
	* the accession number will be preserved.			
FEATURES	Location/Qualifiers			
source	1..153675			
	/organism="Oryza sativa (japonica cultivar-group)"			
	/mol_type="genomic DNA"			
	/cultivar="Nipponbare"			
	/db_xref="taxon:39947"			
	/chromosome="9"			
	/clone="P0646B07"			
ORIGIN				
Query Match	70.7%;	Score 21.2;	DB 2;	Length 153675;
Best Local Similarity	88.5%;	Pred. No. 2.2e+02;		
Matches	22;	Conservative	0;	Mismatches 3; Indels 0; Gaps 0;
5	CTTGTTCAACCGTTCGCTTGTTCCA 30			
Db	19380 CATGTTAACCGTTCGCTTATTCCA 19405			

RESULT 15	
AP004118	
LOCUS	AP004118 155146 bp DNA linear PLN 21-JAN-2004
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2, BAC clone:OJ1148_D05.
ACCESSION	AP004118
VERSION	AP004118.4 GI:41053257
KEYWORDS	
SOURCE	Oryza sativa (japonica cultivar-group)
ORGANISM	Oryza sativa (japonica cultivar-group)
REFERENCE	1 Sasaki, T., Matsumoto, T. and Yamamoto, K. Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC clone:OJ1148_D05 Published Only in Database (2001)
REFERENCE	2 Sasaki, T., Matsumoto, T. and Yamamoto, K. Direct Submission Submitted (29-AUG-2001) Takuji Sasaki, National Institute of Agricultural Sciences, Rice Genome Research Program; Kamondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) On Jan 21, 2004 this sequence version replaced gi:32490504. Genes were predicted from the integrated results of the following: GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), FGENESH (http://www.softberry.com/), GeneMark (http://opal.biology.gatech.edu/GeneMark/), Glimmer (http://www.tigr.org/cdb/glimmer/glmr_form.html), RiceHMM (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4 (http://globin.cse.psu.edu/html/docs/sim4.html), gap2 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The genomic sequence was searched against NCBI NonRedundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP or DBJ. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DBJ accession no. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown protein'. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to IRISP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence. The orientation of the sequence is from -21M13 to M13rev of the BAC clone. This sequence of OJ1148_D05 clone has an overlap with OJ1743_B12 (DBJ: AP004096) clone at 5' end and with P0017H11 (DBJ: AP004786) at 3' end. The sequence was generated by combining Monasato and RGP-Japan sequencing data. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES	Location/Qualifiers
source	1. 155146 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="2" /clone="OJ1148_D05" 5441..9517 /gene="OJ1148_D05.1"
gene	
gene	
gene	join(5441..5906,6423..6676,6786..6894,7393..7531,7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,9165..9517) /gene="OJ1148_D05.1" /note="supported by full-length cDNA(s) : AK105916" join(5463..5906,6423..6676,6786..6894,7393..7531,7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,9165..9517) /gene="OJ1148_D05.1" /note="supported by full-length cDNA(s) : AK101568" join(5577..5906,6423..6676,6786..6894,7393..7531,7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,9165..9274) /gene="OJ1148_D05.1" /note="contains EST(s) : AU076051(E2768), AU076052(E2768) contains full-length cDNA(s) : AK105916, AK101568" /codon_start=1 /product="putative silvertleaf whitefly-induced protein 1" /protein_id="BAD07684.1" /db_xref="GI:41053258" /translation="MASPSLRDVAVGGLDRDPVSLIGKLGISRRRLQNDPPALVPQED LVAOHVVDALPLVSMETGGGLVVRKVSVAEGRSNIVYPTGPERAISFGMHNDV VVANPEMDPDPFSLTFSEDDKLRGRTDCLGHVLAQMLRGLGVKPKGSV IAVPTANENSLITGIGVDGLVGDGLDKNGPLFWIDTADKOPCIGGVTHLTK AIGKLPHSLAKKAIINAMLNEMALKEIOTMYNDPPEHKEKLYKATPSTTKPKM STPGGLNIPESCTISGDIRLPLPSTTSVVKQDEYVHDINENLEKLPTRPVSKY VLEDELNRLGRLEITIDEDITMGVACNLSESGQALCKATBEIYGVHEPYSIGSLPI RELDEGPFVQTRGVGLITLYHAKNEYCLFSDMAQGFVFLSIISLEADV" 9758..9946 /gene="OJ1148_D05.2" <9758..9946 /gene="OJ1148_D05.2" /note="start and end point are not identified" 9758..9946 /gene="OJ1148_D05.2" /note="predicted by Glimmer etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD07685.1" /db_xref="GI:41053259" /translation="MRAORGAARBITSPSTFGROADDEGAGCGCARARPLHFWPS TARRTAREITGPLYRS" complement(10461..10670) /gene="OJ1148_D05.3" complement(<10461..>10670) /gene="OJ1148_D05.3" /note="start and end point are not identified" complement(10461..10670) /gene="OJ1148_D05.3" /note="predicted by Glimmer etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD07686.1" /db_xref="GI:41053260" /translation="MPTIRPROPTLPQDLSLGRARVPSADEBEKPKDGKKRGQV GSRRMWLGRNRGRGNEERTARR" 11212..11718 /gene="OJ1148_D05.4" <11212..>11718 /gene="OJ1148_D05.4" /note="start and end point are not identified" 11212..11718 /gene="OJ1148_D05.4" /note="predicted by FGENESH etc." /codon_start=1 /product="hypothetical protein" /protein_id="BAD07687.1" /db_xref="GI:41053261" /translation="MTPRRRPDHDGLTRSRAGDDPAMPPTGGRAVAARHSRLRV PLAVVELLATRTIGGPTKIRSDSDLTAADGRGPRMGGRGHPALPSRSSSTPRARL SPCIVRHLMRREBRKQGRQRMWADBPFWALIEGYHNFSGLGLELYLVVDSFE BPIGGILQ" complement(join(15074..15178,17827..17877,17931..18029,


```
misc_feature      18062..18190,18223..18322,18729..18861,18967..19111)
                  /gene="OJ1148_D05.5"
                  complement(join(15074..15178,17827..17877,17931..18029,
18062..18190,18223..18322,18729..18861,18967..19111))
                  /gene="OJ1148_D05.5"
                  /note="hypothetical ORF
                  predicted by GENSCAN
                  this category is not included in IRGSP standard"
gene              complement(25292..25586)
                  /gene="OJ1148_D05.6"
                  complement(join(<25292..25412,25471..>25586))
mRNA              /gene="OJ1148_D05.6"
                  /note="start and end point are not identified"
                  complement(join(25292..25412,25471..25586))
CDS               /gene="OJ1148_D05.6"
                  /note="predicted by GeneMark.hmm etc."
                  /codon_start=1
                  /product="hypothetical protein"
                  /protein_id="BAD07688.1"
                  /db_xref="GI:41053262"
                  /translation="MAHGSAGLKAATVTGDMARQRTGREALMRRRRRDCRRRVG
DGAATATRHDAARQWQSEATATAPVRAGLEAG"
gene              complement(join(26131..26319,28092..28139))
misc_feature      /gene="OJ1148_D05.7"
                  complement(join(26131..26319,28092..28139))
                  /gene="OJ1148_D05.7"
                  /note="hypothetical ORF
                  predicted by GENSCAN
                  this category is not included in IRGSP standard"
gene              30668..30991
                  /gene="OJ1148_D05.8"
                  <30668..>30991
mRNA              /gene="OJ1148_D05.8"
                  /note="start and end point are not identified"
                  30668..30991
CDS               /gene="OJ1148_D05.8"
                  /note="predicted by FGENSEH etc."
                  /codon_start=1
                  /product="hypothetical protein"
                  /protein_id="BAD07689.1"
                  /db_xref="GI:41053263"
                  /translation="MPCTSPSCATTCRTPGNHHHGTSELRAASRRASLWEENVLP
YNFPPLSPSPYFPSSSCQHHLAVMGVEVANVTKVVGEGRKIKREEDDEWLNGSHA
ILDIT"
gene              32784..35862
                  /gene="OJ1148_D05.9"
                  join(32784..33234,33666..34250,35333..35862)
mRNA              /gene="OJ1148_D05.9"
                  /note="supported by full-length cDNA(s) : AK060695"
                  join(32872..33234,33666..34250,35333..35506)
CDS               /gene="OJ1148_D05.9"
                  /note="contains full-length cDNA(s) : AK060695"
                  /codon_start=1
                  /product="putative basic-helix-loop-helix transcription
                  factor"
                  /protein_id="BAD07690.1"

Query Match      70.7%; Score 21.2; DB 8; Length 155146;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Search completed: September 16, 2005, 03:01:38
Job time : 102.667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 14.5242 Seconds

(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctctgtcaacgctcgtctgtccca 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	AAH76335	Aah76335 Z. mays M
2	30	100.0	255	AAH76340	Aah76340 Z. mays M
3	30	100.0	1394	AAH76340	Aah76340 Z. mays M
4	30	100.0	1394	AAH76340	Aah76340 Z. mays M
5	30	100.0	1394	AAH76332	Aah76332 Z. mays M
6	30	100.0	1394	AAH76333	Aah76333 Z. mays M
7	24.2	80.7	158	AAH76334	Aah76334 Z. mays M
8	22	73.3	1304	ABO75351	Abq75351 Human Lun
9	20.4	68.0	1542	ABZ13819	Abz13819 Arabidops
10	20.2	67.7	2397	ACA44013	Acad44013 Prokaryot
11	20	66.7	2000	ADJ40785	Adj40785 Plant CDN
12	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
13	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
14	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
15	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
16	19.6	65.3	2000	ADA71994	Ada71994 Rice gene
17	19.4	64.7	2240	AAQ79668	Aaq79668 Sequence
18	19.4	64.7	2243	AAQ79668	Aaq79668 Sequence
19	19.4	64.7	2243	ACD26214	Acad26214 cDNA enco
20	19.4	64.7	349980	ABQ81845	Abq81845 Bifidobac

C 21	19.2	64.0	2000	ADA73340	Ada73340 Rice gene
C 22	19	63.3	1324	ADA71777	Ada71777 Rice gene
C 23	19	63.3	1388	AAQ43356	Aaq43356 Arabidops
C 24	19	63.3	1821	ADBE93879	Adbe93879 Ixodes sc
C 25	19	63.3	5671	AAK90190	Aak90190 Human dig
C 26	19	63.3	5671	ABA20915	Abaz0915 Human ner
C 27	19	63.3	5671	AA639832	Aa639832 Genomic s
C 28	19	63.3	5671	ADB32792	Adb32792 Human nov
C 29	19	63.3	198522	ACN44010	Acn44010 Human gen
C 30	18.8	62.7	883	ABK65288	Abk65288 Arabidops
C 31	18.8	62.7	883	ADD30938	Add30938 Plant yle
C 32	18.8	62.7	883	ADBE37244	Adbe37244 Plant yle
C 33	18.8	62.7	883	AD141918	Ad141918 Plant tra
C 34	18.8	62.7	1083	AA553644	Aa553644 Helicobac
C 35	18.8	62.7	1083	ACA34769	Ac34769 Prokaryot
C 36	18.8	62.7	1899	ABX06821	Abx06821 S. pneumo
C 37	18.8	62.7	1902	AA555583	Aa555583 Streptoco
C 38	18.8	62.7	1902	ACA49918	Ac49918 Prokaryot
C 39	18.8	62.7	1908	ADJ63835	Adj63835 Plant 1p
C 40	18.8	62.7	2000	ADJ41535	Adj41535 Plant CDN
C 41	18.8	62.7	2247	AD57809	Ad57809 Bacterial
C 42	18.8	62.7	2920	ADM02356	Adm02356 Human CDN
C 43	18.8	62.7	3339	ACF79841	Acf79841 Mouse ani
C 44	18.8	62.7	7812	AA61238	Aa61238 DNA enco
C 45	18.8	62.7	9541	AAV52265	Aav52265 Streptoco

ALIGNMENTS

RESULT 1	AAH76335	standard; DNA; 30 BP.
ID	AAH76335	
XX	AAH76335;	
AC		
XX		
DT	29-OCT-2001	(first entry)
XX		
DE	Z. mays Ms45 male tissue-preferred regulatory region fragment.	
XX		
KW	Ms45; male tissue; regulatory region; transcription; male fertility;	
XX	hybrid seed; de.	
OS	Zea mays.	
XX		
PN	W0200160997-A2.	
XX		
PD	23-AUG-2001.	
XX		
PF	13-FEB-2001; 2001WO-US004527.	
XX		
PR	15-FEB-2000; 2000US-00504487.	
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;	
XX		
DR	WPI; 2001-514772/56.	
XX		
PT	A male tissue-preferred regulatory region comprising nucleotide sequences	
XX	essential for initiating transcription of the Ms45 gene useful for	
PT	mediating fertility in a male plant.	
XX		
PS	Claim 5; Page 47; 50pp; English.	
XX		
CC	The invention provides a male tissue-preferred regulatory region (I)	
CC	comprising nucleotide sequences essential for initiating transcription of	
CC	the Ms45 gene. A method of mediating male fertility in a plant is	
CC	provided that involves introducing an expression vector comprising a	
CC	promoter operably linked to (I) into a plant where the exogenous gene	
CC	impacts male fertility of the plant and (II) controls expression of the	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The	
CC	present sequence represents a DNA fragment -152 to -101 bases upstream of	

CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 XX
 SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30

RESULT 2

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

XX AAH76340;
 AC

DT 29-OCT-2001 (first entry)

DE Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;
 KM hybrid seed; promoter; de.
 XX

OS Zea mays.

PN WO200160997-A2.

PD 23-AUG-2001.

PF 13-FEB-2001; 2001WO-US004527.

PR 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

PA Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 PI WPI; 2001-514772/56.

DR WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

CC The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment

XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 39 CATGCTTGTTCACCGTTCGTTGTTCCA 68

RESULT 3

AAK07408
 ID AAK07408 standard; DNA; 1394 BP.
 XX

AC AAK07408;
 XX
 DT 08-JUN-1999 (first entry)

DE Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
 XX

OS Zea mays.

PN WO9859061-A1.

PD 30-DEC-1998.

PF 19-JUN-1998; 98WO-US012895.

PR 23-JUN-1997; 97US-00880499.

XX (PION-) PIONEER HI-BRED INT INC.

PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 DR WPI; 1999-105628/09.

XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region

PT - useful in mediating plant fertility, especially hybrid seed production.
 PT

PS Claim 2; Page 22-23; 39pp; English.

CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolb and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants

XX SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCACCGTTCGTTGTTCCA 30
 |||
 DB 1179 CATGCTTGTTCACCGTTCGTTGTTCCA 1208

RESULT 4

AAK07409
 ID AAK07409 standard; DNA; 1394 BP.

AC AAK07409;
 XX

DT 08-JUN-1999 (first entry)

DE Zea mays Ms45 male tissue-preferred regulatory region.

XX Ms45; male; tissue-preferred; regulatory region; plant cells;
 KM plant tissue; differentiated; hybrid seed; fertility; ss.
 XX

OS Zea mays.

PN WO9859061-A1.

PD 30-DEC-1998.

PF 19-JUN-1998; 98WO-US012895.

PR 23-JUN-1997; 97US-00880499.


```
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue-preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX |||
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 5
XX AAH76332
XX ID AAH76332 standard; DNA; 1394 BP.
XX
XX AAH76332;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX |||
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 7
```

```
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX |||
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 6
XX AAH76333
XX ID AAH76333 standard; DNA; 1394 BP.
XX
XX AAH76333;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX |||
XX 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX
XX RESULT 7
```


AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mayas Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; db.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (i) into a plant where the exogenous gene
CC impacts male fertility of the plant and (ii) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mayas Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX
Query Match 80.7%; Score 24.2; DB 5; Length 158;
Best Local Similarity 89.7%; Pred. No. 1.5;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTGTGCTGTTC 29
DB 25 CATGCTGTTCACCGTGTGCTGTTC 53
RESULT 8
ABQ75351/C
ID ABQ75351 standard; cDNA; 1304 BP.
XX
AC ABQ75351;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human lung specific nucleic acid sequence SEQ ID NO:90.
XX
KM Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
KM LSP; cytosaratic; gene therapy; vaccine; metastasis; lung cancer;
XX squamous cell carcinoma; gene; ss.
OS Homo sapiens.
XX
PN WO200264788-A2.

XX
PD 22-AUG-2002.
XX
PF 20-NOV-2001; 2001WO-US045080.
XX
PR 20-NOV-2000; 2000US-0252054P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;
XX
DR WPI; 2002-657601/70.
XX
PT New lung specific nucleic acid useful in gene therapy or as vaccines for
PT treating lung cancer (e.g. squamous cell carcinoma) or non-cancerous lung
PT diseases, as well as for diagnosing, monitoring or staging these
PT diseases.
XX
PS Claim 1; Page 206; 282pp; English.
XX
XX The present invention describes an isolated lung specific nucleic acid
CC (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
CC sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
CC (b) comprises any of 135 sequences comprising 148 - 3193 base pairs (bp);
CC given in ABQ75262 to ABQ75376; (c) selectively hybridises to (a) or (b);
CC or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
CC protein (LSP) sequences have cytosaratic activity and can be used in gene
CC therapy and vaccines. LSNA and LSPs are useful for diagnosing and
CC monitoring the presence and metastases of lung cancer in a patient. An
CC antibody that specifically binds to an LSP can be used for determining
CC the presence of an LSP in a sample, as well as for treating a patient
CC with lung cancer. Particularly by inducing an immune response against the
CC lung cancer cell expressing the LSNA or LSPs. In particular, these LSNA
CC and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC cancerous disease states in lung
XX
SQ Sequence 1304 BP; 352 A; 347 C; 314 G; 291 T; 0 U; 0 Other;
XX
Query Match 73.3%; Score 22; DB 6; Length 1304;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTGTGCTGTTC 30
DB 1155 CTTTCTTGTTCCTCCGCTTCTGTTC 1126
RESULT 9
ABZ13819
ID ABZ13819 standard; DNA; 1542 BP.
XX
AC ABZ13819;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana scress regulated gene SEQ ID NO 1624.
XX
KM Arabidopsis thaliana; plant; gene; stress; transgenic; db.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
XX
PR 26-JAN-2001; 2001US-0264647P.
XX
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper JF, Kreps J, Wang X, Zhu T;
 XX
 DR WPI; 2002-304127/34.
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 1624; 577bp + Sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid to a
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 SQ Sequence 1542 BP; 345 A; 382 C; 370 G; 445 T; 0 U; 0 Other;
 Query Match 68.0%; Score 20.4; DB 6; Length 1542;
 Best Local Similarity 80.0%; Pred. No. 81;
 Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CATGCTGTTCACCGTTCGCTTGTGCA 30
 Db 237 CATGATTGTTCACCAATCGCTGCTTCCA 266
 RESULT 10
 ID ACMA44013 standard; DNA; 2397 BP.
 AC ACA44013;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #25670.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Pseudomonas putida.
 XX
 PN WO200277183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362659P.
 XX
 PA (ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 XX
 DR P-PSDB; AB040143.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX
 PS Claim 14; SEQ ID NO 31883; 1766bp; English.
 CC
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 SQ Sequence 2397 BP; 415 A; 825 C; 781 G; 376 T; 0 U; 0 Other;
 Query Match 67.3%; Score 20.2; DB 8; Length 2397;
 Best Local Similarity 88.0%; Pred. No. 1e+02;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CATGCTGTTCACCGTTCGCTTG 25
 Db 516 CATGCTGTTCACCGTTCGCTTG 540
 RESULT 11
 ID ADJ40785 standard; cDNA; 2000 BP.
 AC ADJ40785;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Plant cDNA #1785.
 XX
 KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
 KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
 KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
 KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
 KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
 KW antifungal.
 XX
 OS Eukaryota.
 XX
 PN US2004016025-A1.
 XX
 PD 22-JAN-2004.
 XX
 PF 26-SEP-2002; 2002US-00260238.
 XX
 PR 26-SEP-2001; 2001US-0325277P.
 PR 26-SEP-2001; 2001US-0325448P.

PR		04-APR-2002; 2002US-0370620P.
XX	(BUDW//)	BUDWORTH P.
PA	(MOUG//)	MOUGHAMER T.
PA	(BRIG//)	BRIIGS S P.
PA	(COOP//)	COOPER B.
PA	(GLAZ//)	GLAZEBROOK J.
PA	(GOFF//)	GOFF S A.
PA	(KATR//)	KATAGIRI F.
PA	(KEEP//)	KEEPS J.
PA	(PROV//)	PROVART N.
PA	(RICK//)	RICKE D.
XX	(ZHUT//)	ZHU T.
PI	Budworth P,	Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI	Goff SA, Katagiri F,	Kreps J, Provart N, Rieke D, Zhu T;
DR	WPI; 2004-190374/18.	
XX		
PT	New rice promoter; useful for manipulating crop plants to alter or	
PT	improve phenotypic characteristics, e.g. produce large quantities of oil	
PT	or proteins, resistance to insecticides, virus or fungi, stress tolerance	
PT	or high nutritional value.	
XX		
XX	Claim 26; SEQ ID NO 1785; 230bp; English.	
CC	The invention relates to plant nucleotide sequences that direct seed-,	
CC	leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential	
CC	or constitutive transcription of an operatively linked nucleic acid	
CC	segment. The invention also relates to a method for augmenting a plant	
CC	genome and a method of identifying a gene, where its expression is	
CC	altered in the seed, leaf, stem, panicle, pollen, root or is constitutive	
CC	in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,	
CC	cannola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,	
CC	sorghum, rice or wheat. The polynucleotides and the polypeptides they	
CC	encode are useful for manipulating crop plants to alter or improve	
CC	phenotypic characteristics, to produce large quantities of oil or	
CC	proteins, to incur resistance to insecticides, viruses or fungi, and to	
CC	incur stress tolerance (e.g. salt, cold or drought) to ensure the plants	
CC	have a high nutritional value with reduced apical dominance or dwarfism,	
CC	early flowering or altered metabolic pathways. This sequence represents a	
CC	plant nucleic acid of the invention. Note: The sequence data for this	
CC	patent did not form part of the printed specification but was obtained in	
CC	electronic format directly from USPTO at seqdata.uspto.gov/sequence.html .	
XX		
SQ	Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;	
	Query Match	66.7%; Score 20; DB 12; Length 2000;
	Best Local Similarity	82.1%; Pred.No. 1.2e+02;
	Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0	
Oy	3 TGCTTGTTCAACCGTTCGCTTGTTCCA 30	
Ddb	1179 TACATGTTTAACCGTTCGCTTATTCAA 1206	
RESULT 12		
ID	ADA72061	standard; DNA; 2000 BP.
XX	ADA72061;	
AC	ADA72061;	
XX		
DT	20-NOV-2003	(first entry)
XX		
DE	Rice gene, SEQ ID 5386.	
XX		
XX	Plant; bacterial infection; fungal infection; viral infection; rice;	
KM	gene; ds.	
XX		
OS	Oryza sativa.	
XX		
PN	WO2003000898-A1.	
XX		

PD	03-JAN-2003.
XX	
PF	22-JUN-2001; 2001WO-IB001105.
XX	
PR	22-JUN-2001; 2001WO-IB001105.
XX	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI	Katagiri F, Qian S, Tao Y, Whiteham S, Xie Z, Zhu T, Zou G;
XX	
DR	WPI; 2003-175290/17.
XX	
PT	Identifying at least one gene involved in plant resistance or response to
PT	pathogenic infection for conferring resistance or tolerance to a plant to
PT	bacterial, fungal or viral infection by determining or detecting plant
PT	gene expression.
PS	Claim 27; SEQ ID NO 5386; 899pp; English.
XX	
CC	The present invention relates to a method (M1) for identifying genes
CC	involved in plant resistance or response to pathogenic infection. M1
CC	comprises identifying a gene whose expression is significantly altered in
CC	the incompatible interaction of plant gene expression relative to
CC	expression of the gene in an uninfected plant, in a mutant plant that
CC	does not express a gene associated with response to pathogenic infection,
CC	or in a corresponding incompatible or compatible interaction. (M1) is
CC	useful for conferring resistance to resistance or tolerance to a plant to
CC	bacterial, fungal or viral infection. The present sequence was used to
CC	illustrate the invention.
SQ	Sequence 2000 BP; 605 A; 377 C; 326 G; 690 T; 0 U; 2 Other;
Query Match	65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity	84.6%; Pred.No.1.8e+02;
Matches 22; Conservative	0; Mismatches 4; Indels 0; Gaps 0
OY	5 CTGTGTCACCGTTCGCTCTGTGTCGA 30
Db	974 CATGTTTACCGTTCGCTCTATTTCGA 999
RESULT 13	
ADA71994/C	
ID	ADA71994 standard; DNA; 2000 BP.
XX	
AC	ADA71994;
XX	
DT	20-NOV-2003 (first entry)
XX	
DE	Rice gene, SEQ ID 5319.
XX	
KW	Plant; bacterial infection; fungal infection; viral infection; rice;
KW	gene; ds.
XX	
OS	Oryza sativa.
XX	
PN	WO2003000898-A1.
XX	
PD	03-JAN-2003.
XX	
PF	22-JUN-2001; 2001WO-IB001105.
XX	
PR	22-JUN-2001; 2001WO-IB001105.
XX	
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI	Katagiri F, Qian S, Tao Y, Whiteham S, Xie Z, Zhu T, Zou G;
XX	
DR	WPI; 2003-175290/17.
XX	
PT	Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
XX
XX Claim 27; SEQ ID NO 5319; 899pp; English.
PS
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 638 A; 400 C; 347 G; 615 T; 0 U; 0 Other;
SQ
XX
XX Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CTGTTCACCGTTCGTTGTTCCA 30
DB 150 CATGTTAACCGTTCGTTATTCCA 125
RESULT 14
ADA71736/c
ID ADA71736 standard; DNA; 2000 BP.
XX
XX ADA71736;
DT 20-NOV-2003 (first entry)
XX
XX Rice gene, SEQ ID 5061.
DE
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
XX *Oryza sativa*.
OS
XX WO2003000898-A1.
PN
XX
XX 03-JAN-2003.
PD
XX
XX 22-JUN-2001; 2001WO-IB001105.
PF
XX
XX 22-JUN-2001; 2001WO-IB001105.
PR
XX
XX Claim 27; SEQ ID NO 5061; 899pp; English.
PS
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI
XX
XX WPI; 2003-175290/17.
DR
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5061; 899pp; English.
PS
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX

CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 601 A; 422 C; 306 G; 671 T; 0 U; 0 Other;
SQ
XX
XX Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CTGTTCACCGTTCGTTGTTCCA 30
DB 291 CATGTTAACCGTTCGTTATTCCA 266
RESULT 15
ADA72530/c
ID ADA72530 standard; DNA; 2000 BP.
XX
XX ADA72530;
AC
XX
XX 20-NOV-2003 (first entry)
DT
XX
XX Rice gene, SEQ ID 5855.
DE
XX
XX Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
XX *Oryza sativa*.
OS
XX WO2003000898-A1.
PN
XX
XX 03-JAN-2003.
PD
XX
XX 22-JUN-2001; 2001WO-IB001105.
PF
XX
XX 22-JUN-2001; 2001WO-IB001105.
PR
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI
XX
XX WPI; 2003-175290/17.
DR
XX
XX Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
XX Claim 27; SEQ ID NO 5855; 899pp; English.
PS
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
XX Sequence 2000 BP; 549 A; 413 C; 443 G; 588 T; 0 U; 7 Other;
SQ
XX
XX Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 CTGTTCACCGTTCGTTGTTCCA 30
DB 27 CATGTTAACCGTTCGTTATTCCA 2

Search completed: September 15, 2005, 21:33:22
Job time : 15.5242 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 14, 2005, 19:14:19 ; Search time 4.53786 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtctaacgcgtctgtctgtcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing filter 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/5A COMB.seq: *
4: /cgn2_6/prodata/1/ina/5B COMB.seq: *
5: /cgn2_6/prodata/1/ina/5A COMB.seq: *
6: /cgn2_6/prodata/1/ina/5B COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	1394	3	US-08-880-499-1
2	30	100.0	1394	3	US-08-880-499-2
3	19.4	64.7	2243	3	US-07-937-609-15
4	19.4	64.7	2243	3	US-08-029-170-15
5	19.4	64.7	2243	4	US-09-443-745-15
6	19.4	64.7	2243	4	US-09-443-745-15
7	18.8	62.7	8756	3	US-09-368-590-1
8	18.8	62.7	8756	4	US-09-368-590-1
9	18.8	62.7	8756	3	US-09-368-590-1
10	18.8	62.7	8756	4	US-09-368-590-1
11	18.6	62.0	549	4	US-09-949-016-13480
12	18.6	62.0	549	4	US-09-949-016-13480
13	18.6	62.0	549	4	US-09-949-016-13480
14	18.6	62.0	549	4	US-09-949-016-13480
15	18.6	62.0	549	4	US-09-949-016-13480
16	18.6	62.0	549	4	US-09-949-016-13480
17	18.6	62.0	549	4	US-09-949-016-13480
18	18.6	62.0	549	4	US-09-949-016-13480
19	18.6	62.0	549	4	US-09-949-016-13480
20	18.6	62.0	549	4	US-09-949-016-13480
21	18.6	62.0	549	4	US-09-949-016-13480
22	18.6	62.0	549	4	US-09-949-016-13480
23	18.6	62.0	549	4	US-09-949-016-13480
24	18.6	62.0	549	4	US-09-949-016-13480
25	18.6	62.0	549	4	US-09-949-016-13480
26	18.6	62.0	549	4	US-09-949-016-13480
27	18.6	62.0	549	4	US-09-949-016-13480

c 28	17.8	59.3	601	4	US-09-949-016-204030	Sequence 204030,
c 29	17.8	59.3	978	3	US-08-858-207A-174	Sequence 174, App
c 30	17.8	59.3	999	3	US-08-858-207A-174	Sequence 193, App
c 31	17.8	59.3	2009	1	US-07-958-222A-1	Sequence 1, Appl1
c 32	17.8	59.3	8073	4	US-09-949-016-12327	Sequence 12327, A
c 33	17.8	59.3	8074	4	US-09-949-016-16645	Sequence 16645, A
c 34	17.8	59.3	9472	1	US-08-325-547-9	Sequence 9, Appl1
c 35	17.8	59.3	40505	4	US-09-949-016-13439	Sequence 13439, A
c 36	17.8	59.3	72843	4	US-09-949-016-12574	Sequence 12574, A
c 37	17.8	59.3	14115	4	US-09-949-016-17490	Sequence 17490, A
c 38	17.8	59.3	183202	4	US-09-949-016-13614	Sequence 13614, A
c 39	17.8	59.3	265038	4	US-09-949-016-15779	Sequence 15779, A
c 40	17.8	59.3	276237	4	US-09-949-016-17504	Sequence 17504, A
c 41	17.4	58.0	331	4	US-09-640-211A-169	Sequence 169, App
c 42	17.4	58.0	601	4	US-09-949-016-178197	Sequence 178197, App
c 43	17.4	58.0	719	4	US-09-107-433-820	Sequence 820, App
c 44	17.4	58.0	789	4	US-09-252-991A-11034	Sequence 11034, A
c 45	17.4	58.0	813	4	US-09-252-991A-5455	Sequence 5455, Ap

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Fox, Tim W.

APPLICANT: Albertson, Marc C.

APPLICANT: Carl, Garnet W.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 30; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-08-029-170-15

Query Match
Best Local Similarity 64.7%; Score 19.4; DB 3; Length 2243;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy
Db 1 CATGCTGTTCACCGTCTGCTGTTC 29
5 CCTGCTGCTCAACTCTACGCTTGTTC 33

RESULT 5
US-09-443-745-15
Sequence 15, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
FILE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
```

```

APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-09-443-745-15

Query Match
Best Local Similarity 64.7%; Score 19.4; DB 4; Length 2243;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy
Db 1 CATGCTGTTCACCGTCTGCTGTTC 29
5 CCTGCTGCTCAACTCTACGCTTGTTC 33
```

```

RESULT 6
US-09-949-016-16262
Sequence 16262, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CLO01307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16262
LENGTH: 45571
TYPE: DNA
ORGANISM: Human
US-09-949-016-16262

Query Match
Best Local Similarity 63.3%; Score 19; DB 4; Length 45571;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```


Qy	3	TGCTTGTTCAACCGTTGCTTGTTCC	29
Db	38119	TGTTTGTTCAATCGTCTTCTTCTTCC	38145

```

RESULT 7
US-09-368-590-1/c
? Sequence 1, Application US/09368590
? Patent No. 6187563
? GENERAL INFORMATION:
? APPLICANT: Solimena, Michele
? TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
? TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
? FILE REFERENCE: 101918-200 (OCR-941)
? CURRENT APPLICATION NUMBER: US/09368, 590
? CURRENT FILING DATE: 1999-08-04
? EARLIER APPLICATION NUMBER: 60/095, 657
? EARLIER FILING DATE: 1998-08-07
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: FASTSEQ for Windows Version 3.0
? SEQ ID NO 1
? LENGTH: 7812
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(6879)
? NAME/KEY: unsure
? LOCATION: (100)...(102)
? NAME/KEY: unsure
? LOCATION: (1021)...(1023)
? NAME/KEY: unsure
? LOCATION: (2266)...(2268)
US-09-368-590-1

```

Query Match	Similarity	Score	DB	Length
Beet	Local	76.7%	11e+02	7812
Matches	23	Conservative	0	Mismatches 7, Indels 0, Gaps 0.

Qy	1	CATGCTTGTCACCGTCGCTTGTTC	CA	30
Ob	841	CATGCTTCTGCACAGTCGCTGCCTCA	812	

```

RESULT 8
US-09-949-016-1438/c
; Sequence 1438, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1438
; LENGTH: 8756
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1438

```

Query Match	62.7%	Score 18.8;	DB 4;	Length 8756;
Best Local Similarity	76.7%;	Pred. No. 1.2e+02;		
Matches 23;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTCTGTGCCA 30
 ||||| | | | | | | | | |
Db 1912 CATGCTTCTGCACAGAGTGTGTGCCATCA 1883

```

1 RESULT 9
2 US-08-961-527-132
3 / Sequence 132, Application US/08961527
4 / Patent No. 6420135
5 / GENERAL INFORMATION:
6 / APPLICANT: Charles Kunach
7 / TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
8 / NUMBER OF SEQUENCES: 391
9 / CORRESPONDENCE ADDRESS:
10 / ADDRESSEE: Human Genome Sciences, Inc.
11 / STREET: 9410 Key West Avenue
12 / CITY: Rockville
13 / STATE: Maryland
14 / COUNTRY: USA
15 / ZIP: 20850
16 / COMPUTER READABLE FORM:
17 / MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
18 / COMPUTER: HP Vectra 486/33
19 / OPERATING SYSTEM: MSDOS version 6.2
20 / SOFTWARE: ASCII Text
21 / CURRENT APPLICATION DATA:
22 / APPLICATION NUMBER: US/08/961,527
23 / FILING DATE:
24 / CLASSIFICATION: 424
25 / PRIOR APPLICATION DATA:
26 / APPLICATION NUMBER:
27 / FILING DATE:
28 / ATTORNEY/AGENT INFORMATION:
29 / NAME: Brooke, A. Anders
30 / REGISTRATION NUMBER: 36,373
31 / REFERENCE/DOCKET NUMBER: PB340P1
32 / TELECOMMUNICATION INFORMATION:
33 / TELEPHONE: (301) 309-8504
34 / TELEFAX: (301) 309-8512
35 / INFORMATION FOR SEQ ID NO: 132:
36 / SEQUENCE CHARACTERISTICS:
37 / LENGTH: 9541 base pairs
38 / TYPE: nucleic acid
39 / STRANDEDNESS: double
40 / TOPOLOGY: linear
41 US-08-961-527-132

```

Query	Match	Similarity	Score	DB 3	Length
db	Best Local	76.7%	Pred. No. 1.2e+02		
	Matches	23	Conservative	7	Indels 0; Gaps 0
QY	1	CAGCGTTGTCACCGTTCCGCTTGTTC	CA	30	
	8049	CAGCGTTGTCACCGTTCCATTTTTTTC	CA	8078	

```

RESULT 10
US-09-949-016-13180/C
; Sequence 13180, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001037
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/0241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
;

```



```
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: PasteSeq for Windows Version 4.0
/ SEQ ID NO 13180
/ LENGTH: 112874
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(112874)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13180

Query Match          62.7%; Score 18.8; DB 4; Length 112874;
Best Local Similarity 76.7%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCTGTTTCACCGTTCGTTGTTCCA 30
DB 41167 CATGCTTTCGACACAGTGTGCTGCCCA 41138

RESULT 11
US-09-270-767-29480/c
/ Sequence 29480, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 29480
/ LENGTH: 549
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-29480

Query Match          62.0%; Score 18.6; DB 4; Length 549;
Best Local Similarity 84.0%; Pred. No. 84;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGTTCGTTGTT 27
DB 223 TGCTTGTGAAGTTCCTTCTGTT 199

RESULT 12
US-09-270-767-13492/c
/ Sequence 13492, Application US/09270767
/ Patent No. 6703491
/ GENERAL INFORMATION:
/ APPLICANT: Homburger et al.
/ TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
/ FILE REFERENCE: File Reference: 7326-094
/ CURRENT APPLICATION NUMBER: US/09/270,767
/ CURRENT FILING DATE: 1999-03-17
/ NUMBER OF SEQ ID NOS: 62517
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 13492
/ LENGTH: 1120
/ TYPE: DNA
/ ORGANISM: Drosophila melanogaster
US-09-270-767-13492

Query Match          62.0%; Score 18.6; DB 4; Length 1120;
Best Local Similarity 84.0%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGTTCGTTGTT 27
DB 223 TGCTTGTGAAGTTCCTTCTGTT 199
```

```
RESULT 13
US-09-248-796A-3293
/ Sequence 3293, Application US/09248796A
/ Patent No. 6747137
/ GENERAL INFORMATION:
/ APPLICANT: Keith Weinstein et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
/ FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.132
/ CURRENT APPLICATION NUMBER: US/09/248,796A
/ CURRENT FILING DATE: 1999-02-12
/ PRIOR APPLICATION NUMBER: US 60/074,725
/ PRIOR FILING DATE: 1998-02-13
/ PRIOR APPLICATION NUMBER: US 60/096,409
/ PRIOR FILING DATE: 1998-08-13
/ NUMBER OF SEQ ID NOS: 28208
/ SEQ ID NO 3293
/ LENGTH: 1260
/ TYPE: DNA
/ ORGANISM: Candida albicans
US-09-248-796A-3293

Query Match          61.3%; Score 18.4; DB 4; Length 1260;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGTTCGTTGTTCCA 30
DB 903 TGCTTGTTCACGTTGTTGTTCCGCCA 930

RESULT 14
US-09-573-080A-7/c
/ Sequence 7, Application US/09573080A
/ Patent No. 6828097
/ GENERAL INFORMATION:
/ APPLICANT: JOAN, KNOUL
/ APPLICANT: ROGAN, PETER
/ TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI
/ FILE REFERENCE: 30307
/ CURRENT APPLICATION NUMBER: US/09/573,080A
/ CURRENT FILING DATE: 2000-05-16
/ NUMBER OF SEQ ID NOS: 479
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 7
/ LENGTH: 2710
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: repeat region
/ LOCATION: (1)..(2710)
/ OTHER INFORMATION: charlie3
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: n is a, c, g or t
/ PUBLICATION INFORMATION:
/ AUTHORS: Jurka, J; Malchukiewicz, J; Milosavljevic, A
/ TITLE: Prototypic sequences for human repetitive DNA
/ JOURNAL: Journal of Molecular Evolution
/ VOLUME: 35
/ ISSUE: 4
/ PAGES: 286-291
/ DATE: 1992-10-
/ DATABASE ACCESSION NUMBER: Database for repetitive elements (repbase)
/ DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-7

Query Match          60.7%; Score 18.2; DB 4; Length 2710;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```


Qy 3 TGCTGTCAACCGTTCGTTGTT 27
||| ||| ||| ||| ||| ||| |||
Db 1209 TGCTGTTCACGAGTTCGTTKGT 1185

RESULT 15
US-09-949-016-200252/c
; Sequence 200252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200252
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-200252

Query Match 60.0%; Score 18; DB 4; Length 601;
Best Local Similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 5 CTTGTTCAACCGTTCGTTGTTCCA 30
||| ||| ||| ||| ||| ||| |||
Db 239 CTTGTTAAACAGTTCCTGTGCTCCA 214

Search completed: September 15, 2005, 08:25:22
Job time : 6.53786 secs

Sat Sep 17 09:06:48 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 19.5179 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaacgcgttcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10J_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	US-10-713-381-4	Sequence 4, Appli
2	30	100.0	25	US-10-713-381-9	Sequence 9, Appli
3	30	100.0	1394	US-10-713-381-1	Sequence 1, Appli
4	30	100.0	1394	US-10-713-381-2	Sequence 2, Appli
5	24.2	80.7	158	US-10-713-381-3	Sequence 3, Appli
6	22	73.3	1304	US-10-001-857-90	Sequence 90, Appli
7	21	70.0	1980090	US-10-713-993-6615	Sequence 6615, Ap

Result No.	Score	Query Match	Length	ID	Description
8	21	70.0	1980090	US-10-741-600-17676	Sequence 17676, A
9	20.4	68.0	1542	US-09-938-842A-1624	Sequence 1624, Ap
10	20.4	68.0	1542	US-09-938-842A-1624	Sequence 1624, Ap
11	20.2	67.3	2397	US-10-282-122A-31883	Sequence 31883, A
12	20	66.7	2000	US-10-260-238-1785	Sequence 1785, Ap
13	19.6	65.3	640	US-10-437-963-96639	Sequence 96639, A
14	19.6	65.3	901	US-10-425-115-3534	Sequence 3534, Ap
15	19.6	65.3	2000	US-10-260-228-1977	Sequence 1977, Ap
16	19.4	64.7	2243	US-09-443-745-15	Sequence 15, Appli
17	19.4	64.7	2256646	US-10-470-565-1	Sequence 1, Appli
18	19	63.3	1041	US-10-425-115-100318	Sequence 100318, A
19	19	63.3	1821	US-10-425-563-5	Sequence 5, Appli
20	19	63.3	1821	US-10-972-789A-5	Sequence 5, Appli
21	19	63.3	5671	US-09-764-872-729	Sequence 729, App
22	19	63.3	198522	US-10-087-132-244	Sequence 244, App
23	18.8	62.7	883	US-09-934-455-279	Sequence 279, App
24	18.8	62.7	883	US-10-225-068-183	Sequence 183, App
25	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
26	18.8	62.7	883	US-10-374-780A-381	Sequence 381, App
27	18.8	62.7	883	US-10-225-068-183	Sequence 183, App
28	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
29	18.8	62.7	1083	US-09-815-242-7281	Sequence 7281, Ap
30	18.8	62.7	1083	US-10-282-122A-22639	Sequence 22639, A
31	18.8	62.7	1083	US-10-335-977-2194	Sequence 2194, Ap
32	18.8	62.7	1125	US-10-335-977-2195	Sequence 2195, Ap
33	18.8	62.7	1856	US-10-437-963-40048	Sequence 40048, A
34	18.8	62.7	1859	US-10-472-928-2217	Sequence 2217, Ap
35	18.8	62.7	1902	US-09-815-242-9220	Sequence 9220, Ap
36	18.8	62.7	1902	US-10-282-122A-37788	Sequence 37788, A
37	18.8	62.7	2000	US-10-260-238-2535	Sequence 2535, A
38	18.8	62.7	2247	US-10-369-493-33483	Sequence 33483, A
39	18.8	62.7	2220	US-10-108-260A-1041	Sequence 1041, A
40	18.8	62.7	8788	US-10-756-149-64	Sequence 64, Appli
41	18.8	62.7	9441	US-08-961-527-132	Sequence 132, App
42	18.8	62.7	9541	US-10-158-844-132	Sequence 132, App
43	18.6	62.7	2165598	US-10-472-928-4979	Sequence 4979, Ap
44	18.6	62.0	375	US-09-974-300-3242	Sequence 3242, Ap
45	18.4	61.3	174	US-09-814-353-5230	Sequence 5230, Ap

ALIGNMENTS

RESULT 1
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4
Query Match 100.0%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 CATGCTGTTCACCGTCTGTTCCTTCCCA 30

RESULT 2
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 30; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTCTGTTCCTTCCCA 30
Db 39 CATGCTGTTCACCGTCTGTTCCTTCCCA 68

RESULT 3
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTCTGTTCCTTCCCA 30
Db 1179 CATGCTGTTCACCGTCTGTTCCTTCCCA 1208

RESULT 4

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTCTGTTCCTTCCCA 30
Db 1179 CATGCTGTTCACCGTCTGTTCCTTCCCA 1208

RESULT 5
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 80.7%; Score 24.2; DB 20; Length 158;
Best Local Similarity 89.7%; Pred. No. 2.2;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTCTGTTCCTTCCCA 29
Db 25 CATGCTGTTCACCGTCTGTTCCTTCCCA 53

RESULT 6
US-10-001-857-90/c
; Sequence 90, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto


```
/ APPLICANT: Recipon, Hervé
/ APPLICANT: Chen, Sei-Yu
/ APPLICANT: Sun, Yongming
/ APPLICANT: Liu, Chengshua
/ TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
/ FILE REFERENCE: DEX-0273
/ CURRENT APPLICATION NUMBER: US/10/001,857
/ PRIOR APPLICATION NUMBER: 2001-11-20
/ PRIOR FILING DATE: 2000-11-20
/ NUMBER OF SEQ ID NOS: 208
/ SOFTWARE: PatentIn Version 3.1
/ SEQ ID NO 90
/ LENGTH: 1304
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-001-857-90

Query Match      73.3%; Score 22; DB 13; Length 1304;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db      1155 CTTCTTGTTCCTCCCGTCTCTCTGTTCCA 1126

RESULT 7
US-10-719-993-6815/c
/ Sequence 6815, Application US/10719993
/ Publication No. US2004025849A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001496
/ CURRENT APPLICATION NUMBER: US/10/719,993
/ CURRENT FILING DATE: 2003-11-24
/ NUMBER OF SEQ ID NOS: 55342
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 6815
/ LENGTH: 1980090
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(1980090)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match      70.0%; Score 21; DB 20; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 ATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db      1948918 ATTCTTGCTGAACAGTTGCTGTGTACCA 1948890

RESULT 8
US-10-741-600-17676/c
/ Sequence 17676, Application US/10741600
/ Publication No. US2005026169A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ FILE REFERENCE: CL001499
/ CURRENT APPLICATION NUMBER: US/10/741,600
/ CURRENT FILING DATE: 2003-12-22
/ NUMBER OF SEQ ID NOS: 73997
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 17676
```

```
/ LENGTH: 1980090
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(1980090)
/ OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

Query Match      70.0%; Score 21; DB 21; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2 ATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db      1948918 ATTCTTGCTGAACAGTTGCTGTGTACCA 1948890

RESULT 9
US-09-938-842A-1624
/ Sequence 1624, Application US/09938842A
/ Patent No. US20020160378A1
/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ FILE REFERENCE: SCRIPI300-3
/ CURRENT APPLICATION NUMBER: US/09/938,842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300,111
/ NUMBER OF SEQ ID NOS: 5379
/ SEQ ID NO 1624
/ LENGTH: 1542
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match      68.0%; Score 20.4; DB 9; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
Db      237 CATGATTTGTTCAACCAATGTCGTTTCCA 266

RESULT 10
US-09-938-842A-1624
/ Sequence 1624, Application US/09938842A
/ Publication No. US20040009476A9
/ GENERAL INFORMATION:
/ APPLICANT: Harper, Jeff
/ APPLICANT: Kreps, Joel
/ APPLICANT: Wang, Xun
/ APPLICANT: Zhu, Tong
/ TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
/ FILE REFERENCE: SCRIPI300-3
/ CURRENT APPLICATION NUMBER: US/09/938,842A
/ CURRENT FILING DATE: 2001-08-24
/ PRIOR APPLICATION NUMBER: US 60/227,866
/ PRIOR FILING DATE: 2000-08-24
/ PRIOR APPLICATION NUMBER: US 60/264,647
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/300,111
```


;; PRIOR FILING DATE: 2001-06-22
;; NUMBER OF SEQ ID NOS: 5379
;; SEQ ID NO 1624
;; LENGTH: 1542
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624

Query Match 68.0%; Score 20.4; DB 11; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
DB 237 CATGATTGTTCACCAATCGTGGTTTCCA 266

RESULT 11
US-10-282-122A-31883
; Sequence 31883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITTA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent version 3.1
; SEQ ID NO 31883
; LENGTH: 2397
; TYPE: DNA
; ORGANISM: Pseudomonas putida
US-10-282-122A-31883

Query Match 67.3%; Score 20.2; DB 17; Length 2397;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTTG 25
|||||

DB 516 CATGCTGTTCACCGTTCGCTTG 540

RESULT 12
US-10-260-238-1785
; Sequence 1785, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Puniyaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Rieke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1785
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-260-238-1785

Query Match 66.7%; Score 20; DB 17; Length 2000;
Best Local Similarity 82.1%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 TCCTTGTTCACCGTTCGTTGTTCCA 30
DB 1179 TACATGTTTACCGTTCGTTATTCAA 1206

RESULT 13
US-10-437-963-96639/C
; Sequence 96639, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 96639
; LENGTH: 640
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94716C.1
US-10-437-963-96639

Query Match 65.3%; Score 19.6; DB 19; Length 640;

Best Local Similarity 84.6%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CTGTTCACCGTTCGTCCTGTTCGA 30
DB 199 CATGTTAACCGTTCGTCCTATTCAA 174

RESULT 14

US-10-425-115-3534/c
Sequence 3534, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 3534
LENGTH: 901
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_103225C.1
US-10-425-115-3534

Query Match 65.3%; Score 19.6; DB 20; Length 901;
Best Local Similarity 84.6%; Pred. No. 2.5e+02;

Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACCGTTCGTCCTGTGT 26
DB 33 CACGCTGTTCACCGTTCCTCGGTGT 8

RESULT 15

US-10-260-238-1977/c
Sequence 1977, Application US/10260238
Publication No. US20040016025A1
GENERAL INFORMATION:
APPLICANT: Budworth, Paul R.
APPLICANT: Moushmer, Todd G.
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazedbrook, Jane
APPLICANT: Golf, Stephen A.
APPLICANT: Katagiri, Fumiyuki
APPLICANT: Kreps, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Rieke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 60111-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/325,448
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/325,277
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 6077
SEQ ID NO 1977
LENGTH: 2000
TYPE: DNA
ORGANISM: Oryza sativa
US-10-260-238-1977

Query Match 65.3%; Score 19.6; DB 17; Length 2000;
Best Local Similarity 84.6%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 CTGTTCACCGTTCGTCCTGTTCGA 30
DB 291 CATGTTAACCGTTCGTCCTATTCAA 266

Search completed: September 15, 2005, 20:46:05
Job time : 26.5379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 93.5658 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcacacgcgtcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	30	100.0	687	9	CC656939 OGMDO20TV
C 2	30	100.0	915	9	CG224225 OG1AG08TV
C 3	30	100.0	963	9	CC656933 OGMDO20TV
C 4	22.2	74.0	308	9	CG700474 ZMMB0C017
C 5	21.2	70.7	534	5	BU035739 CHJ7N19.Y
C 6	21.2	70.7	626	5	BU030489 CHJ15120
C 7	21.2	70.7	792	7	CN137013 OX1_54.B0
C 8	21	70.0	124	1	AA811551 0A99603.S
C 9	21	70.0	351	9	CR192459 Reverse s
C 10	21	70.0	571	9	CR064270 Reverse s
C 11	21	70.0	599	9	CR233964 Reverse s
C 12	21	70.0	769	9	CR188179 Reverse s
C 13	21	70.0	832	9	CC848693 NDL_2114.
C 14	21	70.0	835	9	CC860012 NDL_2113.
C 15	20.6	68.7	412	7	R12648
C 16	20.6	68.7	450	7	R12618
C 17	20.6	68.7	582	1	AU290776
C 18	20.6	68.7	681	8	AZ817790
C 19	20.6	68.7	1144	8	CC271930 CH261-112
C 20	20.4	68.0	353	8	AQ261553 CITBI-ET1
C 21	20.4	68.0	356	2	BE522930 M30D4STM
C 22	20.4	68.0	514	4	BG354154 00661_1ea
C 23	20.4	68.0	516	5	BP561510 BP561510
C 24	20.4	68.0	536	1	AV439701 AV439701

C 25	20.4	68.0	559	1	AV825937
C 26	20.4	68.0	560	6	CA230882
C 27	20.4	68.0	600	8	BH876120
C 28	20.4	68.0	602	9	CL583422 OB_Ba007
C 29	20.4	68.0	603	8	BZ694195 SP_Ba004
C 30	20.4	68.0	615	8	AQ453944 HS_5165.B
C 31	20.4	68.0	668	8	AO260121 CITBI-ET1
C 32	20.4	68.0	779	6	CD825116
C 33	20.4	68.0	807	6	CD836362
C 34	20.4	68.0	836	8	BH480110
C 35	20.4	68.0	967	8	BZ694358
C 36	20.4	68.0	1205	8	B10570
C 37	20.2	67.3	431	7	R32995
C 38	20.2	67.3	565	6	CA681397
C 39	20.2	67.3	729	8	BZ052097
C 40	20.2	67.3	791	9	CC836492
C 41	20	66.7	170	9	DR5E15T
C 42	20	66.7	227	4	B1130928
C 43	20	66.7	321	4	B1127375
C 44	20	66.7	329	7	CK098862
C 45	20	66.7	368	4	B1128587

ALIGNMENTS

RESULT 1
LOCUS CC656939/c
DEFINITION OGMDO20TV ZM 0.7 1.5 KB Zea mays genomic clone ZMM8MA0554D15,
genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE

AUTHORS Whiteclaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
COMMENT Other GSSs: OGMDO20TM
Contact: Cathy Whiteclaw
TIGR

TITLE

JOURNAL 9712 Medical Center Drive, Rockville, MD 20850, USA
COMMENT Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteclaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES

source Location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMM8MA0554D15"
/clone_1lb="ZM_0.7 1.5_KB"
/note="Vector: pBCSK-; Site: 1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CATGCTGTTCACACGCTCGTCTGTTCAC 30
|||||
397 CATGCTGTTCACACGCTCGTCTGTTCAC 368

RESULT 2
 CG224225 915 bp DNA linear GSS 22-AUG-2003
 LOCUS
 DEFINITION OG1AG08TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0716B15,
 genomic survey sequence.
 ACCESSION CG224225 GI:34124113
 VERSION
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 915)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OG1AG08TV
 Contact: Cathy WhiteIaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 912 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteIaw@tigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 1..915
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0716B15"
 /note="Vector: pBCSk-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 100.0%; Score 30; DB 9; Length 915;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 Db 558 CATGCTTGTTCAACCGTTCGTTGTTCCA 587

RESULT 3
 CC656933 963 bp DNA linear GSS 19-JUN-2003
 LOCUS
 DEFINITION OGMDQ20TM ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0554D15,
 genomic survey sequence.
 ACCESSION CC656933
 VERSION
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 963)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other_GSSs: OGMDQ20TV
 Contact: Cathy WhiteIaw
 TIGR

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteIaw@tigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 1..963
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMA0554D15"
 /note="Vector: pBCSk-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 100.0%; Score 30; DB 9; Length 963;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 Db 776 CATGCTTGTTCAACCGTTCGTTGTTCCA 805

RESULT 4
 CG700474 308 bp DNA linear GSS 15-OCT-2003
 LOCUS
 DEFINITION ZMMBMC0174H15f ZMMBMC (BCORI) Zea mays genomic clone ZMMBMC0174H15
 5', genomic survey sequence.
 ACCESSION CG700474
 VERSION
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 308)
 Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
 Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
 Sequencing of the maize genome at PGIR (2003c)
 Unpublished (2003)
 Contact: Bharti,A.K.
 Dr. Joachim Messing's lab
 The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
 University
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
 Tel: 732 445 3801
 Fax: 732 445 5735
 Email: bharti@waksman.rutgers.edu
 Seq primer: T7
 Class: BAC ends
 High quality sequence start: 57.
 Location/Qualifiers
 1..308
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBMC0174H15"
 /lab_host="B. coli DH10B"
 /clone_lib="ZMMBMC (BCORI)"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI"

ORIGIN
 Query Match 74.0%; Score 22.2; DB 9; Length 308;
 Best Local Similarity 88.9%; Pred. No. 49;
 Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTT 27

DB	226	CATGCTGGTTACGCCGTTCTTCTT	200
RESULT 5	BU035739	534 bp	mRNA linear EST 23-AUG-2002
LOCUS	OHJTN19.YG.ab1	OH_EFGHJ sunflower	RHA280 Helianthus annuus cDNA
DEFINITION	clone OHJTN19, mRNA sequence.		
ACCESSION	BU035739		
VERSION	BU035739.1	GI:22471259	
KEYWORDS	EST.		
SOURCE	Helianthus annuus (common sunflower)		
ORGANISM	Helianthus annuus		
REFERENCE	Eukaryota; Vitidiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.		
AUTHORS	1 (bases 1 to 534) Kozlik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Reeseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, U., Ellison, P., Kolman, J., Slabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.		
JOURNAL	Leucute and Sunflower ESTs from the Compositae Genome Project http://compgenomics.ucdavis.edu/ unpublished (2002)		
COMMENT	Contact: Alexander Kozlik [R.W.Michelmore] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Ammundson Hall, UCD, Davis, CA 95616, USA Tel: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozlik@ucdavis.edu [michelmore@vegmail.ucdavis.edu] belongs to contig OH_CA_Conctis3046, see http://cgsdb.ucdavis.edu/ for details. Plate: OHJ7 row: N column: 19.		
FEATURES	Location/Qualifiers		
source	1..534 /organism="Helianthus annuus" /mol_type="mRNA" /cultivar="RHA280" /db_xref="taxon:4232" /clone="OHJTN19" /lab_host="E.coli" /clone_1lib="OH_EFGHJ sunflower RHA280" /ncseq="Vector: pBRCDNASFIAB. The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgsdb.ucdavis.edu/ TAG_TISSUE=flowers environmental stress TAG_LIB=OH_EFGHJ sunflower RHA280 TAG_SEQ=CGATGCGCG"		
ORIGIN			
Query Match	70.7%	Score 21.2;	DB 5; Length 534;
Best Local Similarity	88.5%	Pred. No. 1.5e+02;	
Matches	23; Conservative	0; Mismatches	3; Indels
			0; Gaps
0y	2 ATGCTTGTCAACGTCGTCCTTGT 27		
Db	104 ATGCTTGTCAACGTCGTCCTTGT 79		
RESULT 6	BU030489	626 bp	mRNA linear EST 23-AUG-2002
LOCUS	BU030489/c		
DEFINITION	OHJ15120.YG.ab1	OH_EFGHJ sunflower	RHA280 Helianthus annuus cDNA
	clone OHJ15120, mRNA sequence.		

ACCESSION	BUI030489
VERSION	BUI030489.1
KEYWORDS	GI:22466009
SOURCE	EST.
ORGANISM	Helianthus annuus (common sunflower)
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Helianthus.
AUTHORS	Kozik A., Michelmore R.W., Knapp S., Chevalier M., Riesberg L., Lin H., van Damme M., Lavelle D., Shattler P., Ziegler J., Ellison P., Kolman J., Slabaugh M.S., Livingston K., Zhou Y., Lettice C., Church S., Jackson L. and Bradford K. http://compugenomics.ucdavis.edu/ Unpublished (2002)
JOURNAL	Contact: Alexander Kozik [R.W.Michelmore@vegmail.ucdavis.edu] Department of Vegetable Crops, R.W.Michelmore Lab University of California at Davis (UCD) Aasmundson Hall, UCD, Davis, CA 95616, USA Tel.: 1-(530)-742-1742 Fax: 1-(530)-752-9659 Email: akozik@ucg.org [michelmore@vegmail.ucdavis.edu] belongs to contig OH_Ca_contig3046, see http://cgdb.ucdavis.edu/ for details.
FEATURES	Plate: OHJ15 row: I column: 20.
SOURCE	Location/Qualifiers 1..626 /organism="Helianthus annuus" /mol_type="mRNA" /cultivar="RHA280" /db_xref="taxon:4232" /clone="OHJ15120" /lab_host="E.coli" /cloned_lib="OH_ERGHU sunflower RHA280" /note="Vector: pBKCDSAS1Lib: The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformation details made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/TAG_TISSUE=flowers environmental stress TAG_Lib=OH_ERGHU sunflower RHA280 TAG_SEQ=CGRATCGGAG"
ORIGIN	
Query Match	70.7%; Score 21.2; DB 5; Length 626;
Best Local Similarity	88.5%; Pred. No. 1.5e+02;
Matches	23; Conservative 0; Mismatches 3; Indels 0; Gaps 0,
Dy	2 ATGCTGTTCACCGTTCGTTGTT 27 DB 104 ATGCTTCTCAAGCCTTCGTTT 79
RESULT 7	
LOCUS	CNI37013 792 bp mRNA linear EST 01-APR-2004
DEFINITION	bicolor cDNA clone OX1_54_B01_A002 5', mRNA sequence.
ACCESSION	CNI37013
VERSION	CNI37013.1
KEYWORDS	GI:45970259
SOURCE	Sorghum bicolor (sorghum)
ORGANISM	Sorghum bicolor Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 792)

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 351;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTTGTCACCGCTTCGTGTTCCA 30
DB 307 ATGCTGTCACCGCTTCGTGATGTTCCA 279

RESULT 10

CR064270 571 bp DNA linear GSS 05-JUL-2004
LOCUS CR064270/c
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN190119, genomic survey sequence.
ACCESSION CR064270
VERSION CR064270.1 GI:49797742
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
1 (bases 1 to 571)
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES
source
1..571
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN190119"
/clone_1lb="MHPN"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 571;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTTGTCACCGCTTCGTGTTCCA 30
DB 309 ATGCTGTCACCGCTTCGTGATGTTCCA 281

RESULT 11
CR233964 599 bp DNA linear GSS 06-JUL-2004
LOCUS CR233964/c
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN144h08, genomic survey sequence.
ACCESSION CR233964
VERSION CR233964.1 GI:50012813
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 599)
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES
source
1..599
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN144h08"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 599;
Best Local Similarity 82.8%; Pred. No. 1.8e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTTGTCACCGCTTCGTGTTCCA 30
DB 310 ATGCTGTCACCGCTTCGTGATGTTCCA 282

RESULT 12
CR188179 769 bp DNA linear GSS 06-JUL-2004
LOCUS CR188179/c
DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and
chromosome engineering clone MHPN234h03, genomic survey sequence.
ACCESSION CR188179
VERSION CR188179.1 GI:49967028
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 769)
TITLE Direct Submision
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES
source
1..769
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPN234h03"
/clone_1lb="MHPN"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 769;
Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTTGTCACCGCTTCGTGTTCCA 30
DB 307 ATGCTGTCACCGCTTCGTGATGTTCCA 279

RESULT 13
CC848693 832 bp DNA linear GSS 24-JUL-2003
LOCUS CC848693/c
DEFINITION ND1.2114.T7 Notre Dame Liverpool Aedes aegypti genomic clone
Notredame Liverpool-2114, genomic survey sequence.
ACCESSION CC848693
VERSION CC848693.1 GI:33197857
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti

REFERENCE
AUTHORS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;
Stegomyia.
1 (bases 1 to 832)
TITLE Aedes aegypti
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: ND1.2114.SP6
Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208

Email: enta@ligr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..832
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="Notre Dame Liverpool-2114"
/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 832;

Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY

2 ATGCTTGTTCACCGTTCGTTGTTCCA 30

DB

242 ATCCTTGTTCACCATTTGCTTGATTC 214

RESULT 14
CC860012/c 835 bp DNA linear GSS 24-JUL-2003

LOCUS ND1.2113.T7 Notre Dame Liverpool Aedes aegypti genomic clone

DEFINITION Notre Dame Liverpool-2113, genomic survey sequence.

ACCESSION CC860012

VERSION CC860012.1 GI:33220022

KEYWORDS GSS.

SOURCE Aedes aegypti (yellow fever mosquito)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes;

Stegomyia.

1 (bases 1 to 835)
Loftus, B., Shetty, J., Knudson, D. and Severson, D.

BAC end sequencing of Aedes aegypti

Unpublished (2003)

Other GSS: ND1.2113.SP6

Contact: Brenden Loftus

Department of Eukaryotic Genomics

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-3543

Fax: 301-838-0208

Email: enta@ligr.org

Library was provided by David Severson

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1..835
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="Liverpool"
/db_xref="taxon:7159"
/clone="Notre Dame Liverpool-2113"
/note="Vector: pECBAC1, Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 70.0%; Score 21; DB 9; Length 835;

Best Local Similarity 82.8%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATGCTTGTTCACCGTTCGTTGTTCCA 30
|||
DB 242 ATCCTTGTTCACCATTTGCTTGATTC 214

FEATURES

source

1..412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:398607"
/db_xref="taxon:9606"
/clone="IMAGE:26260"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: Lambda BAC; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
5'-GGGAGGAGGATTCGCGCGGAGGAGGATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BAC vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 68.7%; Score 20.6; DB 7; Length 412;

Best Local Similarity 85.2%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCCTTGTTCACCGTTCGTTGTTGCC 29

DB

313 TACTTGTTCATCTGTTGATTC 339

TITLE The Washu-Merck EST Project
JOURNAL Unpublished (1995)
CONTACT: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1704

FEATURES

source

High quality sequence stops: 276 Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lln.gov) for further information.
Insert Length: 1704 Std Error: 0.00
Seq primer: M13RP1

FEATURES

source

Location/Qualifiers
1..412
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:398607"
/db_xref="taxon:9606"
/clone="IMAGE:26260"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: Lambda BAC; Site 1: Not I; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
5'-GGGAGGAGGATTCGCGCGGAGGAGGATTTTCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lambda BAC vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 68.7%; Score 20.6; DB 7; Length 412;

Best Local Similarity 85.2%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCCTTGTTCACCGTTCGTTGTTGCC 29

DB

313 TACTTGTTCATCTGTTGATTC 339

Search completed: September 16, 2005, 08:08:43
Job time : 93.5658 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 134.222 Seconds
(without alignments)
1440.280 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggaatcctaccctcccaaacatccatcctaccatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_ncg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_srs:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	40	100.0	40	6	AX224339 Sequence
2	40	100.0	50	6	AX224338 Sequence
3	40	100.0	158	6	AX224336 Sequence
4	40	100.0	255	6	AX224402 Sequence
5	40	100.0	1394	6	AX224394 Sequence
6	40	100.0	1394	6	AX224395 Sequence
7	40	100.0	1394	6	BD062176 Male t188
8	40	100.0	1394	6	BD062177 Male t188
9	40	100.0	3343	8	AF360356 Zea mays
10	30.4	76.0	137327	8	AF360356 Zea mays
11	24	60.0	55001	9	AC084290 Oryza sat
12	24	60.0	127196	10	AL807784 Mouse DNA
13	24	59.0	171940	2	AC148514 Papio anu
14	23.6	59.0	193735	2	AC148501 Callithrix
15	23	57.5	117026	9	AC008897 Homo sapi
16	23	57.5	149232	2	AC102478 Mus muscu
17	23	57.5	177104	2	AC125876 Rattus no
18	23	57.5	195919	2	CR847932 Danio rer
19	23	57.5	207208	2	CR788256 Danio rer

20	23	57.5	218224	10	AC134404 Mus muscu
21	23	57.5	234163	2	AC115318 Rattus no
22	23	57.5	242109	2	AC147523 Oryza sat
23	23	57.5	319056	2	AC016640 Homo sapi
24	22.8	57.0	164944	3	AC008195 Drosophil
25	22.8	57.0	175781	3	AC008091 Drosophil
26	22.8	57.0	235928	3	AB003736 Drosophil
27	22.6	56.5	95886	2	AL156238 Human DNA
28	22.6	56.5	130893	2	AC149638 Human DNA
29	22.6	56.5	160246	2	AC009221 Homo sapi
30	22.6	56.5	161817	10	AC132119 Mus muscu
31	22.6	56.5	170973	9	AP000923 Homo sapi
32	22.6	56.5	175110	9	AP002428 Homo sapi
33	22.6	56.5	186457	2	AC129669 Rattus no
34	22.6	56.5	196490	2	AC134013 Rattus no
35	22.6	56.5	235953	2	AC130117 Rattus no
36	22.2	55.5	112361	9	H839B17 Human DNA
37	22.2	55.5	21700	9	BX322645 Human DNA
38	22.2	55.5	71561	2	AL928982_6 Continuation (7 of
39	22.2	55.5	73885	2	AC025307 Homo sapi
40	22.2	55.5	111122	8	ATP11C1 Arabidops
41	22.2	55.5	150489	2	AC069491 Homo sapi
42	22.2	55.5	152797	2	AC016856 Homo sapi
43	22.2	55.5	153814	2	AP003180 Homo sapi
44	22.2	55.5	157574	9	AL592293 Human DNA
45	22.2	55.5	176084	2	AC113354 Homo sapi

ALIGNMENTS

RESULT 1	AX224399	Sequence 6 from Patent WO0160997.	40 bp	DNA	linear	PAT 10-SRP-2001
LOCUS	AX224399					
DEFINITION	AX224399					
ACCESSION	AX224399					
VERSION	AX224399.1	GI:15554641				
KEYWORDS						
SOURCE						
ORGANISM	Zea mays					
REFERENCE	1	Albertsen,M.C., Fox,T.W., Garmaat,C.W., Huffman,G. and Kendall,T.L.				
AUTHORS						
TITLE		Male tissue-preferred regulatory region and method of using same				
JOURNAL		Patent: WO 0160997-A 6 23-AUG-2001;				
PIONEER		HI-BRD INTERNATIONAL, INC. (US)				
FEATURES		Location/Qualifiers				
source		1..40				
ORIGIN		/organism="Zea mays"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:4577"				
Query Match		100.0%; Score 40; DB 6; Length 40;				
Best Local Similarity		100.0%; Pred.No. 4.2e-05;				
Matches	40; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	AGGAATCCTACTCTCCCAACATCCATCTTACTCATGCAAC 40				
	1	AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAAC 40				
Db	1	AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAAC 40				
RESULT 2	AX224398	Sequence 5 from Patent WO0160997.	50 bp	DNA	linear	PAT 10-SRP-2001
LOCUS	AX224398					
DEFINITION	AX224398					
ACCESSION	AX224398					
VERSION	AX224398.1	GI:15554640				
KEYWORDS						
SOURCE						
ORGANISM	Zea mays					

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 5 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

Db 11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

RESULT 3
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX224396.1 GI:15554638
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

Db 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

RESULT 4
AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX224402.1 GI:15554644
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

Db 99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

RESULT 5
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX224394.1 GI:15554636
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

Db 1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

RESULT 6
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
AX224395.1 GI:15554637
KEYWORDS
SOURCE
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS
TITLE
JOURNAL
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 1278

RESULT 7
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062176.1 GI:22607781
ACCESSION BD062176.1
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A-1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 1278

RESULT 8
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177.1 GI:22607782
ACCESSION BD062177.1
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A-2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BD 30-OCT-2001
PF 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PI MAC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAA,T,GARY A HUFFMAN,
TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/10,C12N9/00 PC
C07K14/34,C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 40
|||||
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAC 1278

RESULT 9
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS Zea mays male fertility protein (Ms45) gene, complete cds.
DEFINITION AF360356.1 GI:14028756
ACCESSION AF360356.1
VERSION
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Cloning of Ms45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Ms45"
/join<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="Ms45"
/product="male fertility protein"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRNLOWRRGRDGIVOYPHLFPALALALIVADPGLSPLAEV
DYRPVKEHLAPYGEVWGSWPRDNASRLRGRLEFVGEVGESEIFDLOGRPVAGLA
DGRVVRWNGEAGMETFAVMNDGSEVCANGVNSTRKQHEKEFCGRPIGLRPHGE
TGEILVVAAYGVGLAVVGOSGVASVAREADGDP1RFPANDLDVHRNGSVFPDTSRY
SKKHLNILLBEGEGRLRYDPEISGVHVVLKGLVFPNGVQISDHOPLPSPTNC
RIMRYWLEGPRAGEVVFANLDGFPDNRVNSRGQFPWVAIDCCRTPADEVFAKRWLR

ORIGIN

TLTYKPEPLSLKVLTWKARRMHTVLLLDGEGRVVLEDRGHEWKLVSSEVSEVGRK
LMTGVAHHNIATIPYPLED"

Query Match 100.0%; Score 40; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 2,6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCATCTTACATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACATCATCTTACATGCAAC 1278

RESULT 10
AC135206/c 137327 bp DNA linear PLIN 16-APR-2003
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) chromosome 3 clone
AC135206
VERSION
AC135206.3 GI:27596977
KEYWORDS
HTG.
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and
Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and
Sun,S.
Rice Genomic Sequence
Unpublished
2 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., Collura,K. and Thompson,S.
Direct Submision
Submitted (09-OCT-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
3 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C.,
Currie,J., Collura,K. and Thompson,S.
Direct Submision
Submitted (06-NOV-2002) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
4 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J.,
Collura,K. and Thompson,S.
Direct Submision
Submitted (11-JAN-2003) Arizona Genomics Institute, University of
Arizona, 303 Forbes, Tucson, AZ 85721, USA
5 (bases 1 to 137327)
Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T.,
Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and
Sun,S.
Direct Submision
Submitted (16-APR-2003) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA
On Jan 11, 2003 this sequence version replaced gi:24635891.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., paired quality
>=30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by more than
one plasmid subclone; and the assembly was confirmed by
restriction digest. There are transposons located at area
29427-32874 and at area 19945-33212. Region 70511-70536 is covered
by a single subclone. Region 18678-18673 is a double stranded
single subclone. Areas 69957-70761, 89836-90416, 98951-99009,
105316-105607, and 133052-133152 are covered by Monsanto only. The
nucleotide sequence of this BAC clone was generated by combining
Syngenta, Monsanto and Arizona Genomics Institute sequencing data.
Location/Qualifiers
1. .137327
/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"
/db_xref="taxon:39947"
/chromosome="13"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
complement(638..9495)
/gene="OJ1041F02.1"
/note="unknown protein"

CDS
complement(join(638..838,1141..1276,3132..3244,3326..3421,
3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,
5322..5411,5484..5699,5808..5942,6035..6538,6622..6837,
7548..7593,7679..7815,8196..8328,9221..9495))
/gene="OJ1041F02.1"
/codon_start=1
/product="unknown protein"

/protein_id="AAP06844.1"
/db_xref="GI:29893590"
/translation="MAQSLLELLIQFLMPNDNARQAEQIRRLARDPQVPALVHLL
RTAKTPNVRQLAAVLIRKKTITSHWPKLPFAKASLQKALDISITDHSILVRAGANV
VSIKAVVAPGEMPELPIFOGSGPOEDHREVALIFSLTETGTTGSHLNDL
OPIILKICLODASSSRVRIALKAQGSFIEVNDGSDVVKIRPDPISTILNVSROLAN
GEBDVASTAFETPDELIESPAFLGDSRSTIVQPSLEVCNSDELEINRQOIQIISM
LVKFASTLKHKLVTPIQVWCPLTSTAEDGSDSLAABRSAAEVIDTMAINPRH
VPPPLAEASVSFRHINPKYREAATSLGVVSEGCERHKKEDCLVVEALADQD
QVNRGAAPFALGQFAHEHQPEILSHYSEVLICINALBDSPSEVSEKSYVALAACE
MGENTLPIPLDPLMCRLVMSLQSPRNLOETGMSAIGSAVAAAEQAFMEYAEVLEMK
GPMVLTNDEDCARARATEVGIYMAVGRARMETILPPIEATISGVLDVSELEY
THGFSNVAETLIDSGFAOYLPHVVPALSSCNLDGSDVDDSDSUNGSGVSSDD
DVNDEPRVRIISWLTITADITAIIRAIPPAADVLEKODILDTWNIYIKTRWEDD
KEVVAQACTSLADIVRDCGFALIEPYITRLADATLILRQSCCQOVSDGDDGDDID
HDEVLMADVSDLLPAFAKMGVSYDPDITPKLFDLSMKPAKSPHPQDRTMVATLAEV
AQMGAPISAVYDKIMPLVLKELASBATNRNAAFCVGENCKNGAALAKYGDILH
GLHRLFADSEBDADVADRNAGALIRMIWVQOSIPLNQCILPVPVIVNAPQVAVSPN
ESDEKYTVAKAVASHLISVGOQNPILSALPFAHANAASPARNR"
12372..15503
/gene="OJ1041F02.2"
/note="unknown protein"
join(12372..12757,12858..13025,13130..13295,13716..13826,
13982..14130,14224..14354,14459..14558,15197..15503)
/gene="OJ1041F02.2"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"
/translation="MPQOGECCRHGMLVGEBAAVGGGEPEFYVPLKRLSVDSKAS
TAPRTICEGCEBAGDITCDIYAAPLRSCGSAKAMPPEAPLPRMTPPEPPRQSGE
EARRPGEAIRKGRHSYSMLNLIQGISISVSKSALLPQKLAASFPDRKXWTRFP
EGSKFTPPHSHVDFRWKDYCPAVERPHKLFGVDAEYMLAI CGDITRELASPGKSG
SCFTTQDDRMIKTVKKSSEVKVILRMIRSYEHVROKSTLTLTFYETHCKQMGCT
KIHREFDLKSGSHGRTIDTERKIDETTLDLQVAFRLQRFYEELMQLQMDCT
PLENGIWDYSLILGVHFRNDYSVKISQIHIAEPKSTGRKSPGSGSPCELCFEV
SGCDRODIDSRKPEIOLGIMPPQAEBSKQKILNDPLNRHILPITPSSGSCDYVL
FGGIIIDLIQDYDITKLEHAYKSQVNDYISANDPKLYSRFPDPIRNVFIKEQ"
16734..16878
/note="putative MITE, Gaijin/Gaijin-like"
19263..19454
/note="putative transposon, CACTA, En/Spm-like"
21819..24660
/gene="OJ1041F02.3"
/note="hypothetical protein"
join(21819..22007,22096..22206,23772..23996,24523..24660)
/gene="OJ1041F02.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"
/translation="WATSDSGSGGEGQRLLSIPKEGRIIAPTRPDGTLRKAIRI
RAGYVPEEVAIVYQSGAKQKSGPDVPGYDPLADAPKTKAARNRRREKQOAS
TTNDGKGLHIEDAGETNDXNDVDSYTKQISGIAESILVATSRSDATDNGSGS
SAPDIDKIRALKKKIRALEAAYQVGDPENLKPEOLEKKKIKIGWKEELKLENSSPA
AS"
complement(26563..30575)
/gene="OJ1041F02.4"

FEATURES
source
1. .137327
/organism="Oryza sativa (japonica cultivar-group)"

[illegible]

Db	119478	AGAACCTACTCTCCAAACATCATCGCTGACTCATGCAC	119439
RESULT 11		1 AGAGTACCTACTCTCCAAACATCATCGCTGACTCATGCAC	40
AC084290/c			
LOCUS			
DEFINITION	AC084290	55001 bp DNA	linear
ACCESSION	AC084290	12 BAC RP11-56865 (Roewell Park Cancer Institute Human BAC library) complete sequence.	PRI 28-MAR-2002
VERSION	AC084290.15	GI:19774277	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Cnariata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 55001)		
	Mzuy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooke,S.L., Amarantune,H.C., Aye,J.R., Ayele,M., Banks,T., Barbacia,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burckett,C., Butrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denu,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Bocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhai,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gottrell,U.H., Guvera,W., Gunaratne,P., Hale,S., Hamilton,K., Han,U., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Hollway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulik,S., Hume,J., Ioshihara,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kurehi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Ilea,C., Liu,J., Liu,W., Louissege,H., Lozato,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Mardonel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metcsher,S., Metzner,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Monabdet,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,S., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwonn,G., Ocranugue,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Piams,E., Pull,L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojdukkan,I., Rolfe,M., Ruiz,S., Savery,G., Scherzer,S., Scott,G., Shen,H., Shin,C., Shooshitari,N., Stenson,I., Sodegrem,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svarek,A., Tabor,P., Taneirisa,A., Taneirisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vaguer,L., Vera,V., Villalón,D., Vinson,R., Wang,S., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kuchelapatti,R., Weinstein,G., and Gibbs,R.		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 55001)		
REFERENCE	Worley,K.C.		
AUTHORS	Submitted (21-OCT-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
JOURNAL	3 (bases 1 to 55001)		
REFERENCE	Worley,K.C.		
AUTHORS	Submitted (27-MAR-2002) Human Genome Sequencing Center, Department		

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 55001)

Mortley, K.C.

Direct Submission

Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Mar 28, 2002 this sequence version replaced gi:19747063.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality/info/genbank.annotation.html>.

FEATURES

QUALSTAT-REPORT.

source Location/Qualifiers

1. .55001
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-568G5"
1. .1301
/note="overlaps bases 177088..178388 of clone AC079630"
/function="clone overlap"
35..229
/rpt_family="MIR"
repeat_region
1002..1023
/rpt_family="AT_rich"
1092..1127
/rpt_family="MIR"
repeat_region
complement(2156..2209)
/rpt_family="L2"
repeat_region
3563..3650
/rpt_family="(CCGTA)n"
complement(3688..3765)
/rpt_family="L2"
repeat_region
complement(3840..3908)
/rpt_family="L1MD3"
repeat_region
complement(4234..4464)
/rpt_family="MIR"

repeat_region	complement(4847..4914)
repeat_region	/rpt_family="MIR"
repeat_region	4938..5053
repeat_region	/rpt_family="L2"
repeat_region	6350..6660
repeat_region	/rpt_family="AluSg"
repeat_region	complement(7309..7743)
repeat_region	/rpt_family="LTR43"
repeat_region	7744..9144
repeat_region	/rpt_family="L1PA4"
repeat_region	complement(9151..9317)
repeat_region	/rpt_family="LTR43"
repeat_region	10587..10846
repeat_region	/rpt_family="L1M4"
repeat_region	10912..11015
repeat_region	/rpt_family="L1P"
repeat_region	11098..11271
repeat_region	/rpt_family="L1ME"
repeat_region	11450..11635
repeat_region	/rpt_family="MLT1"
repeat_region	11947..12179
repeat_region	/rpt_family="MIR"
repeat_region	12797..12828
repeat_region	/rpt_family="(CATATA)n"
repeat_region	complement(12975..13087)
repeat_region	/rpt_family="FLAM_C"
repeat_region	13714..13737
repeat_region	/rpt_family="AT_rich"
repeat_region	14058..14219
repeat_region	/rpt_family="MER45"
repeat_region	14411..14545
repeat_region	/rpt_family="MIR"
repeat_region	complement(15058..15127)
repeat_region	/rpt_family="MIR"
repeat_region	16186..16413
repeat_region	/rpt_family="MIR"
repeat_region	complement(17624..17922)
repeat_region	/rpt_family="AluDb"
repeat_region	complement(18309..18606)
repeat_region	/rpt_family="AluSx"
repeat_region	18844..18854
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(18855..19135)
repeat_region	/rpt_family="AluDb"
repeat_region	19136..19158
repeat_region	/rpt_family="AT_rich"
repeat_region	complement(20660..20753)
repeat_region	/rpt_family="MER91C"
repeat_region	22179..22680
repeat_region	/rpt_family="TAAAA)n"
repeat_region	22180..22358
repeat_region	/rpt_family="AluDo"
repeat_region	22358..22632
repeat_region	/rpt_family="AluDb"
repeat_region	22633..22680
repeat_region	/rpt_family="(TAAAA)n"
repeat_region	22682..22789
repeat_region	/rpt_family="AluDo"
repeat_region	22786..23099
repeat_region	/rpt_family="AluY"
repeat_region	complement(23147..23227)
repeat_region	/rpt_family="MIR"

Query Match 60.0%; Score 24; DB 9; Length 55001;
Best Local Similarity 84.4%; Pred. No. 48;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATACCTACTCCCAACATCATCTTACTCAT 35
Db 47598 AAATTACTCCCAACATCATCTTACTCAT 47567

RESULT 12

AL807784/c
 LOCUS AL807784 127196 bp DNA linear ROD 13-NOV-2002
 DEFINITION Mouse DNA sequence from clone RP23-448C18 on chromosome X, complete sequence.
 ACCESSION AL807784
 VERSION AL807784.11 GI:25045332
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 127196)
 Howden, P.
 Direct Submission
 Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Nov 15, 2002 this sequence version replaced gi:24939941.
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-448C18 is from the RGC1-23 Mouse PAC library
 constructed by the group of Pieter de Jong.
 For further details see http://www.choiri.org/bacpac/home.htm
 VECTOR: pBACe3.6.
 Location/Qualifiers
 1..127196
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone="RP23-448C18"
 /clone_1b="RGC1-23"
 ORIGIN
 Query Match 60.0%; Score 24; DB 10; Length 127196;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 1 AGGATACCTACTCTCCAAATCATCTTACTCATGCAAC 40
 Db 110413 AGGATACATTAAGCAATCACTTCTTACTTAATGCAC 110374
 RESULT 13
 AC148514/c
 LOCUS AC148514 171940 bp DNA linear HTG 24-MAR-2004

DEFINITION Papio anubis clone R41-375M17, WORKING DRAFT SEQUENCE, 3 ordered pieces.
 ACCESSION AC148514
 VERSION AC148514.2 GI:45680463
 KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
 SOURCE Papio anubis (olive baboon)
 ORGANISM Papio anubis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Papio.
 1 (bases 1 to 171940)
 Antonellis, A., Ayle, K., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B., Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P., Larson, S., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Marulles, E.H., Masiello, C., Maskeri, B., McDowell, J., Mullikin, J.C., Paguitigan, C., Portnoy, M.E., Prasad, A., Puri, O., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantirpop, S., Thomas, J.W., Thomas, J.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D., Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 171940)
 Green, E.D.
 Direct Submission
 Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 171940)
 Green, E.D.
 Direct Submission
 Submitted (24-MAR-2004) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Mar 24, 2004 this sequence version replaced gi:45332443.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc_zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: ffn
 Center clone name: 375M17

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones; overlaps with neighboring clones; alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap, version 0.990119
 Consensus quality: 171395 bases at least Q40
 Consensus quality: 171621 bases at least Q30
 Consensus quality: 171701 bases at least Q20
 Insert size: 173000; agarose-fp
 Insert size: 171740; sum-of-contigs
 Quality coverage: 8.27x in Q20 bases; agarose-fp
 Quality coverage: 8.33x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have


```

* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 10148: contig of 10148 bp in length
* 10149 10248: gap of unknown length
* 10249 103023: contig of 92775 bp in length
* 103024 103123: gap of unknown length
* 103124 171940: contig of 68817 bp in length.
FEATURES
    source
        1..171940
            /organism="Papio anubis"
            /mol_type="genomic DNA"
            /db_xref="taxon:9555"
            /clone="RP41-375M17"
            /clone_lib="RP41"
            /note="BAC resource: http://bacpac.chori.org/"
            1..10148
                /note="assembly_fragment"
                clone_end:SP6
                vector_side:left"
            10249..103023
                /note="assembly_fragment"
            103124..171940
                /note="assembly_fragment"
                clone_end:T7
                vector_side:right"
ORIGIN
Query Match      60.0%; Score 24; DB 2; Length 171940;
Best Local Similarity 84.4%; Pred. No. 42;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 ATACCTACTCCCAACATCATCTACTCAT 35
DB 33556 AACTTACTCTCCCAACATCATCTACTCAT 33525

RESULT 14
AC148501/c 193735 bp DNA linear HTG 09-JUL-2004
LOCUS Callitrix jacchus clone CH259-368E20, WORKING DRAFT SEQUENCE, 11
DEFINITION Ordered pieces.
ACCESSION AC148501
VERSION AC148501.3 GI:50080339
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Callitrix jacchus (white-tufted-ear marmoset)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
Callitrix.
1 (bases 1 to 193735)
Antoneillis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Boutefard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Dakl,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
Karling,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-O.,
Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
Masello,C., Maskerl,B., Mcowell,J., Mullikin,J.C., Paguirigan,C.,
Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddi-Dugue,N.,
Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantirpop,S.,
Thomas,J.W., Thomas,P.J., Teipouri,V., Vogt,J.L., Wetherby,K.D.,
Young,A. and Green,E.D.
MISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 193735)
JOURNAL
REFERENCE
AUTHORS Green,E.D.
TITLE Direct Submision
JOURNAL Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Groveomrt Circle, Gaithersburg, MD 20877, USA
REFERENCE
AUTHORS 3 (bases 1 to 193735)
TITLE Direct Submision

```

```

JOURNAL Submitted (09-JUL-2004) NIH Intramural Sequencing Center, 8717
Groveomrt Circle, Gaithersburg, MD 20877, USA
On Jul 9, 2004 this sequence version replaced gi:46430810.
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hri.nih.gov
----- Project Information
Center project name: fvw
Center clone name: 368E20

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 190717 bases at least Q40
Consensus quality: 191709 bases at least Q30
Consensus quality: 192323 bases at least Q20
Insert size: 14400; agarose-fp
Insert size: 192735; sum-of-contigs
Quality coverage: 13.33x in Q20 bases; sum-of-contigs
Quality coverage: 9.96x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 29309: contig of 29309 bp in length
29310 29409: gap of unknown length
29410 37712: contig of 8303 bp in length
37713 37812: gap of unknown length
37813 40472: contig of 2660 bp in length
40473 40572: gap of unknown length
40573 45724: contig of 5152 bp in length
45725 45824: gap of unknown length
45825 47844: contig of 2020 bp in length
47845 47944: gap of unknown length
47945 61027: contig of 13083 bp in length
61028 61127: gap of unknown length
61128 63312: contig of 2085 bp in length
63313 63312: gap of unknown length
63313 70560: contig of 7248 bp in length
70561 70660: gap of unknown length
70661 73658: contig of 2998 bp in length
73659 73758: gap of unknown length
73759 103653: contig of 29895 bp in length
103654 103753: gap of unknown length
103754 193735: contig of 89982 bp in length.
FEATURES
    source
        1..193735
            /organism="Callitrix jacchus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9483"
            /clone="CH259-368E20"
            /clone_lib="CH259"

```


misc_feature /note="BAC resource: http://bacpac.chori.org/"
1..29309
/note="assembly_fragment"
clone_end:T7
vector_side:left"
29410..37712
/note="assembly_fragment"
37813..40472
/note="assembly_fragment"
40573..45724
/note="assembly_fragment"
45825..47844
/note="assembly_fragment"
47945..61027
/note="assembly_fragment"
61128..63212
/note="assembly_fragment"
63313..70560
/note="assembly_fragment"
70661..73658
/note="assembly_fragment"
73759..103653
/note="assembly_fragment"
103754..193735
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
127712..193735
/note="clone overlaps with GenBank Accession Number
AC148551 clone CH259-256024 (center project name fvv)"

ORIGIN

Query Match 59.0%; Score 23.6; DB 2; Length 193735;
Best Local Similarity 76.3%; Pred. No. 60;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 GGATACCTACCTCCAAACATCCATCTTACTCATGCAA 39

DB 89109 GGAAATCACTCCAAAATAATCCCTCAATCATGCAA 89072

RESULT 15
AC008897 117026 bp DNA linear PRI 07-NOV-2000
LOCUS Homo sapiens chromosome 5 clone CTD-2235C13, complete sequence.
AC008897
AC008897.7 GI:11119448
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Unpublished
2 (bases 1 to 117026)
DOE Joint Genome Institute.
Direct Submision
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submision
Submitted (07-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 7, 2000 this sequence version replaced gi:7711370.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.5.

STS Content:
WI-9815 G05428
SHGC-16751 G15410
WI-7171 G06431.
Location/Qualifiers
1..117026
Source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2235C13"

ORIGIN

Query Match 57.5%; Score 23; DB 9; Length 117026;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTACCTCCAAACATCCATCTTACTCATGCAA 39

DB 72634 AGGAAACGAGTTAACAACATCCATCTTCTCAAGCAA 72672

Search completed: September 16, 2005, 03:01:40
Job time : 136.222 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 19.3656 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatactactaccctcccaacacatccatctaccatgcagcacc 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_GeneSeq_16Dec04:*
1: genebegn1980a:*
2: genebegn1990a:*
3: genebegn2000a:*
4: genebegn2001a:*
5: genebegn2001b:*
6: genebegn2002a:*
7: genebegn2002b:*
8: genebegn2003a:*
9: genebegn2003b:*
10: genebegn2003c:*
11: genebegn2003d:*
12: genebegn2004a:*
13: genebegn2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	40	5	AAH76337	Aah76337 Z. mays M
2	40	100.0	50	5	AAH76336	Aah76336 Z. mays M
3	40	100.0	158	5	AAH76334	Aah76334 Z. mays M
4	40	100.0	255	5	AAH76340	Aah76340 Z. mays M
5	40	100.0	1394	2	AAH07408	Aax07408 Zea mays
6	40	100.0	1394	2	AAH07409	Aax07409 Zea mays
7	40	100.0	1394	5	AAH76332	Aah76332 Z. mays M
8	40	100.0	1394	5	AAH76333	Aah76333 Z. mays M
9	22.4	56.0	226215	11	ACN45146	Acn45146 Human ger
10	21.6	54.0	4110	8	ABZ10202	Abz10202 Haematopo
11	21.6	54.0	6309	6	ABL23204	Ab132304 Human imm
12	21.6	54.0	7110	10	ADB54282	Adb54282 Pretreat
13	21.6	54.0	7110	10	ADB54282	Adb54282 Pretreat
14	21.6	54.0	7110	13	ADS89580	Ads89580 Oligonuc
15	21.4	53.5	1341	6	ABZ13750	Abz13750 Arabidops
16	21.4	53.5	6681	6	ABL2155	Ab131155 Human imm
17	21.4	53.5	6681	6	ABL54304	Ab154304 Chemical
18	21.4	53.5	8168	6	ABL32928	Ab132928 Human imm
19	21.4	53.5	8168	6	ABL53328	Aas63328 Chemical
20	21.4	53.5	24401	4	ABL03396	Ab103396 Drosophi

C 21	21	52.5	2000	12	ADJ41594	Adj41594 Plant cDN
C 22	21	52.5	2792	12	ADM98941	Adm98941 Diterpene
C 23	21	52.5	3117	12	ADM98942	Adm98942 Diterpene
C 24	21	52.5	5507	8	ABZ10204	Abz10204 Haematopo
C 25	21	52.5	5507	8	ABZ10058	Abz10058 Haematopo
C 26	21	52.5	8404	4	AAS46500	Aas46500 Tumour su
C 27	21	52.5	8404	6	ABL33595	Ab133595 Human imm
C 28	21	52.5	8404	10	ADB54216	Adb54216 Pretreat
C 29	21	52.5	8404	10	ADB54088	Adb54088 Pretreat
C 30	21	52.5	8404	10	ADB84076	Adb84076 Human lym
C 31	21	52.5	8404	10	ADB84152	Adb84152 Human lym
C 32	21	52.5	8404	13	ADS89242	Ads89242 Oligonuc1
C 33	21	52.5	8404	13	ADS89516	Ads89516 Oligonuc1
C 34	21	52.5	10286	4	AAS45308	Aas45308 Chemical1
C 35	21	52.5	10286	6	ABK28147	Abk28147 DNA trans
C 36	21	52.5	63115	10	ADC85419	Adc85419 Mouse Tle
C 37	21	52.5	63294	9	ADA02939	Ada02939 Mouse Tle
C 38	21	52.5	63294	10	ADB72677	Adb72677 Mouse Tle
C 39	21	52.5	63294	12	ADM74534	Adm74534 Murine ca
C 40	21	52.5	106315	11	ACN43966	Acn43966 Human gene
C 41	20.8	891	8	ABZ51893	Abz51893 Aspergill1	
C 42	20.8	52.0	6620	4	AAS45488	Aas45488 Chemical1
C 43	20.8	52.0	6620	6	ABK28416	Abk28416 DNA trans
C 44	20.8	52.0	6620	6	ABN80279	Abn80279 Human che
C 45	20.8	52.0	110000	5	AAI61373_1	Continuation (2 of

ALIGNMENTS

RESULT 1	AAH76337	standard; DNA; 40 BP.
ID	AAH76337	
XX	AAH76337;	
AC		
XX		
DT	29-OCT-2001	(first entry)
XX		
DE	Z. mays Ms45 male tissue-preferred regulatory region fragment.	
XX		
KW	Ms45; male tissue; regulatory region; transcription; male fertility;	
XX	hybrid seed; ds.	
OS	Zea mays.	
XX		
PN	W0200160997-A2.	
XX		
PD	23-AUG-2001.	
XX		
PF	13-FEB-2001; 2001WO-US004527.	
XX		
PR	15-FEB-2000; 2000US-00504487.	
XX		
PA	(PION-) PIONEER HI-BRED INT INC.	
XX		
PI	Albertsen MC, Fox TW, Garnaat CW, Hufman G, Kendall TL;	
XX		
DR	WPI; 2001-514772/56.	
XX		
PT	A male tissue-preferred regulatory region comprising nucleotide sequences	
XX	essential for initiating transcription of the Ms45 gene useful for	
PT	mediating fertility in a male plant.	
XX		
PS	Claim 14; Page 32; 50pp; English.	
XX		
CC	The invention provides a male tissue-preferred regulatory region (I)	
CC	comprising nucleotide sequences essential for initiating transcription of	
CC	the Ms45 gene. A method of mediating male fertility in a plant is	
CC	provided that involves introducing an expression vector comprising a	
CC	promoter operably linked to (I) into a plant where the exogenous gene	
CC	impacts male fertility of the plant and (I) controls expression of the	
CC	exogenous gene. A method of producing hybrid seeds is also provided. The	
CC	present sequence represents a DNA fragment upstream of the TATA box of a	

CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40

RESULT 2
AAH76336
ID AAH76336 standard; DNA; 50 BP.
XX
AC AAH76336;
XX
DT 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 13-FEB-2001; 2001WO-US004527.
XX
XX PR 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX PS Claim 5; Page 47; 50pp; English.
XX
XX CC The invention provides a male tissue-preferred regulatory region (I)
XX CC comprising nucleotide sequences essential for initiating transcription of
XX CC the Ms45 gene. A method of mediating male fertility in a plant is
XX CC provided that involves introducing an expression vector comprising a
XX CC promoter operably linked to (I) into a plant where the exogenous gene
XX CC impacts male fertility of the plant and (I) controls expression of the
XX CC exogenous gene. A method of producing hybrid seeds is also provided. The
XX CC present sequence represents a DNA fragment -72 to -111 bases upstream of
XX CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX CC nucleotide sequence
XX
SQ Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
XX

Query Match 100.0%; Score 40; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 50

RESULT 3
AAH76334
ID AAH76334 standard; DNA; 158 BP.

XX
AC AAH76334;
XX
DT 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 13-FEB-2001; 2001WO-US004527.
XX
XX PR 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX PS Claim 5; Page 47; 50pp; English.
XX
XX CC The invention provides a male tissue-preferred regulatory region (I)
XX CC comprising nucleotide sequences essential for initiating transcription of
XX CC the Ms45 gene. A method of mediating male fertility in a plant is
XX CC provided that involves introducing an expression vector comprising a
XX CC promoter operably linked to (I) into a plant where the exogenous gene
XX CC impacts male fertility of the plant and (I) controls expression of the
XX CC exogenous gene. A method of producing hybrid seeds is also provided. The
XX CC present sequence represents a DNA fragment -38 to -195 bases upstream of
XX CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX

Query Match 100.0%; Score 40; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 125

RESULT 4
AAH76340
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 promoter fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; promoter; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX


```
PF 13-FEB-2001; 2001MO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL,
XX WPI, 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Example 5; Fig 8; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a Z. mays Ms45 promoter fragment
XX
XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 40; DB 5; Length 255;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-06;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 40
XX |||||||
XX DB 99 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 138
XX
XX RESULT 5
XX AAX07408
XX ID AAX07408 standard; DNA; 1394 BP.
XX
XX AAX07408;
XX
XX DT 08-JUN-1999 (first entry)
XX
XX DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
XX KM Ms45; male; tissue-preferred; regulatory region; plant cells;
XX plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
XX OS Zea mays.
XX
XX PN WO9859061-A1.
XX
XX PD 30-DEC-1998.
XX
XX PF 19-JUN-1998; 98MO-US012895.
XX
XX PR 23-JUN-1997; 97US-00880499.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL,
XX WPI, 1999-105628/09.
XX
XX DR MPI, 1999-105628/09.
XX
XX PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX PS Claim 2; Page 22-23; 39pp; English.
XX
XX CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
XX CC region. It may be used in the construction of a vector for a method of
XX CC producing exogenous genes in a male tissue-preferred manner, which is
```

```
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
XX SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 40; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 6.1e-06;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 40
XX |||||||
XX DB 1239 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 1278
XX
XX RESULT 6
XX AAX07409
XX ID AAX07409 standard; DNA; 1394 BP.
XX
XX AAX07409;
XX
XX DT 08-JUN-1999 (first entry)
XX
XX DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
XX KM Ms45; male; tissue-preferred; regulatory region; plant cells;
XX plant tissue; differentiated; hybrid seed; fertility; ss.
XX
XX OS Zea mays.
XX
XX PN WO9859061-A1.
XX
XX PD 30-DEC-1998.
XX
XX PF 19-JUN-1998; 98MO-US012895.
XX
XX PR 23-JUN-1997; 97US-00880499.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL,
XX WPI, 1999-105628/09.
XX
XX DR MPI, 1999-105628/09.
XX
XX PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
XX PS Claim 3; Page 23-24; 39pp; English.
XX
XX CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
XX CC region. It may be used in the construction of a vector for a method of
XX CC producing exogenous genes in a male tissue-preferred manner, which is
XX CC useful in restoring or conferring fertility, such as in hybrid seed
XX CC production. In conferring fertility, a monocot/dicot plant is transformed
XX CC with the exogenous nucleotide sequence (a male sterility gene, preferably
XX CC Ms45), which encodes a product selected from auxins, rolB and diptheria
XX CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX CC and infertile plants
XX
XX SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 40; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 6.1e-06;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 40
XX |||||||
XX DB 1239 AGGATACCTTACTCCCAACATTCATCTTACTCATGCAAC 1278
```



```
RESULT 7
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a nucleic acid sequence encoding an Ms45 male
XX -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 6.1e-06;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 8
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
```

```
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a nucleic acid sequence encoding an Ms45 male
XX -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 6.1e-06;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCCATCTTACTGATGCAAC 1278
XX
RESULT 9
ACN45146
ID ACN45146 standard; DNA; 226215 BP.
XX
AC ACN45146;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1639824.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1948; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
```


are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Query Match 56.0%; Score 22.4; DB 11; Length 226215;
Best Local Similarity 72.5%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAAC 40
183206 AGAATTCCTGCTTCCAAACACAGCATCTTACTCTTTTAC 183245

RESULT 10
AB210202/c
ID AB210202 standard; DNA; 4110 BP.
XX
XX AB210202;
XX
XX 16-JUN-2003 (first entry)
XX
XX Haematopoietic cell proliferation disorder related DNA sequence #342.
XX
XX Human; haematopoietic cell proliferation disorder; cytostatic;
KM Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KM cytosine methylation state; gene; de.
OS Homo sapiens.
XX
XX MO200277272-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002MO-EP003401.
XX
XX 26-MAR-2001; 2001US-0278333P.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Leache R, Leu B;
PI Lewin A, Lipscher B, Maier S, Model F, Mueller V, Otto T, Pellet C;
PI Schwabe I, Ziebarth H;
XX
XX WPI; 2003-018942/01.
XX
XX Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
XX Claim 28; SEQ ID NO 342; 117bp; English.
XX
XX The present invention describes a method for detecting and
XX differentiating between hematopoietic cell proliferative disorders
XX associated with at least 1 gene and/or their regulatory regions in a
XX subject. The method comprises contacting a target nucleic acid in a
XX biological sample obtained from the subject with at least 1 reagent,
XX which distinguishes between methylated and non-methylated CpG
XX dinucleotides within the target nucleic acid. AB209861 to AB211118
XX represent specifically claimed nucleotide sequences from the present
XX invention. Oligonucleotides from the present invention can be used: for

differentiating between healthy haematopoietic cells and proliferative
disorder haematopoietic cells; for differentiating between acute
lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
determining the cytosine methylation state and/or single nucleotide
polymorphisms (SNPs) of haematopoietic cell proliferation disorder
related sequences and their complements; and as primers for the
amplification of haematopoietic cell proliferation disorder related DNA
sequences. The nucleotide sequences from the present invention can also
be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
disorders allowing for improved and informed treatment of patients

Query Match 54.0%; Score 21.6; DB 8; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

4 ATACCTACTCTCCCAACATCCATCTTAC 31
3083 AAACCTAACCCCAACATCCATCTTAC 3056

RESULT 11
ABL32304/c
ID ABL32304 standard; DNA; 6309 BP.
XX
XX ABL32304;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 277.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
KM antiarteriosclerotic; antianaemic; cyrostatic; nocturnal;
KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KM antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KM antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
de.
XX
XX Homo sapiens.
XX
XX OS
XX MO200200928-A2.
XX
XX 03-JUN-2002.
XX
XX 02-JUL-2001; 2001MO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
XX Claim 1; SEQ ID NO 277; 32bp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 XX
 SQ Sequence 6309 BP; 1758 A; 69 C; 1364 G; 3118 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 6; Length 6309;
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;
 Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATTCATCTACTGATGCAA 39
 DB 3335 ATACCTATCTCCATACGACCAATCTCAATATCCAA 3300

RESULT 12
 ADB54282/c
 ID ADB54282 standard; DNA; 7110 BP.

XX ADB54282;

DT 04-DEC-2003 (first entry)

DE Pretreated genomic DNA region 206.

KW colon cell proliferative disorder; non methylated CpG dinucleotide;
 KW cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.

XX Unidentified.

OS

XX WO2003072821-A2.

PD 04-SEP-2003.

PF 27-FEB-2003; 2003WO-EP002035.

PR 27-FEB-2002; 2002EP-00004551.

XX (EPIG-) EPIGENOMICS AG.

PA Adorjan P, Burger M, Maier S, Nimnrich I, Becker E, Lesche R;

PI Rujan T, Schmitt A;

XX MPI; 2003-731620/69.

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 4 ATACCTACTCCCAACATTCATCTTAC 31
 DB 3083 AAACCTAACCCCAACATTCATCTTAC 3056

RESULT 13
 ADB54196/c
 ID ADB54196 standard; DNA; 7110 BP.

XX ADB54196;

DT 29-JAN-2004 (first entry)

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

KW ds; lymphoid cell proliferative disorder; methylation;
 KW methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
 KW diffuse large B-cell lymphoma; mantle cell lymphoma;
 KW chronic lymphocytic leukemia; small lymphocytic lymphoma;
 KW follicular lymphoma; diagnosis; prognosis.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTAC 31
 |||||
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 14
 ADS89580/C
 ID ADS89580 standard; DNA; 7110 BP.
 XX
 AC ADS89580;
 XX
 XX 18-NOV-2004 (first entry)
 DT
 DE Oligonucleotide of the invention SEQ ID NO:596.
 XX
 XX ss; cell proliferative disorder; breast; methylation; cytostatic;
 KM gene therapy; single nucleotide polymorphism; SNP.
 XX
 OS Unidentified.
 XX
 XX MO2004035803-A2.
 PN
 XX 29-APR-2004.
 PD
 XX 01-OCT-2003; 2003WO-EP010881.
 PF
 XX 01-OCT-2002; 2002DE-01045779.
 PR 07-JAN-2003; 2003DE-01000096.
 PR 17-APR-2003; 2003DE-01017955.
 XX
 XX (EPIC-) EPICGENOMICS AG.
 PA
 XX
 XX Fockens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
 PI Mammrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;
 XX
 XX WPI; 2004-348468/32.
 DR
 XX
 PT Predicting responsiveness of a subject with breast cell proliferative
 PT disorder, useful for treating or differentiating breast cell
 PT proliferative disorders comprises analyzing methylation pattern of a
 PT genomic DNA from the subject.
 PS
 PS Disclosure; SEQ ID NO 596; 104bp; English.
 XX
 XX The invention relates to a novel method for predicting the responsiveness
 CC of a subject with a cell proliferative disorder of the breast tissues to
 CC a therapy comprising analysing the methylation pattern of a target
 CC nucleic acid by contacting at least one of the target nucleic acids in a
 CC biological sample obtained from the subject prior to or during treatment.
 CC The method of the invention has cytosolic activity, and may have a use
 CC in gene therapy. The set of oligonucleotides comprising at least two of
 CC the oligomers are useful for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
 CC methods, nucleic acid, oligonucleotide, and kit are useful for the
 CC treatment, characterisation, classification and/or differentiation, of
 CC breast cell proliferative disorders. The method is also useful for
 CC predicting the responsiveness of a subject with a cell proliferative
 CC disorder of the breast tissues to a therapy. The present sequence is used
 CC in the exemplification of the invention.
 CC
 XX
 XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;
 SQ

Query Match 54.0%; Score 21.6; DB 13; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTAC 31
 |||||
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 15
 AB213750/C
 ID AB213750 standard; DNA; 1341 BP.
 XX
 AC AB213750;
 XX
 XX 21-JAN-2003 (first entry)
 DT
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1555.
 XX
 XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 XX MO200216655-A2.
 PN
 XX 28-FEB-2002.
 PD
 XX 24-AUG-2001; 2001WO-US026685.
 PF
 XX 24-AUG-2000; 2000US-0227866P.
 PR 26-JAN-2001; 2001US-0264647P.
 PR 22-JUN-2001; 2001US-0300111P.
 XX
 XX (SCRI) SCRIPPS RES. INST.
 PA (SYGN.) SYNGENTA PARTICIPATIONS AG.
 XX
 XX Harper JF, Kreps J, Wang X, Zhu T;
 PI
 XX WPI; 2002-304127/34.
 DR
 XX
 PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.
 XX
 XX Claim 14; SEQ ID NO 1555; 577bp + Sequence Listing; English.
 PS
 PS The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office
 CC
 XX
 XX Sequence 1341 BP; 460 A; 231 C; 319 G; 331 T; 0 U; 0 Other;
 SQ

Query Match 53.5%; Score 21.4; DB 6; Length 1341;
 Best Local Similarity 71.8%; Pred. No. 1.5e+02;
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGATATCTACTCCCAACATCATCTTACTGACAA 39
 |||||
 DB 288 AGTCCACCTATTCACCAACATCTCTGTGCTCACACAA 250

Search completed: September 15, 2005, 21:33:23
 Job time : 20.3656 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 14, 2005, 19:14:19 ; Search time 6.05048 Seconds

(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggataccaccaccccaacaacacccatccaccatgcagac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile61.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	1394	3	US-08-880-499-1
2	40	100.0	1394	3	US-08-880-499-2
3	20.6	51.5	1071	4	US-09-248-796A-5546
4	20.6	51.5	42075	4	US-09-949-016-14895
5	20.4	51.0	601	4	US-09-949-016-65770
6	20.4	51.0	973	4	US-09-774-528-184
7	20.4	51.0	4201	3	US-08-945-056-4
8	20.4	51.0	51770	4	US-09-949-016-13668
9	20.4	51.0	84571	4	US-09-949-016-17420
10	20.4	51.0	126200	4	US-09-949-016-11824
11	20.2	50.5	5099	3	US-09-610-040-5
12	20.2	50.5	5099	3	US-10-267-763-5
13	20.2	50.5	5099	3	US-09-949-016-88655
14	20	50.0	601	4	US-09-620-312D-22
15	20	50.0	2615	4	US-09-949-016-13220
16	20	50.0	77586	4	US-09-949-016-13221
17	20	50.0	77586	4	US-09-949-016-13221
18	19.8	49.5	107	4	US-09-513-999C-19002
19	19.8	49.5	601	4	US-09-710-279-2867
20	19.8	49.5	1017	4	US-09-949-016-202698
21	19.8	49.5	1164	3	US-09-134-001C-2534
22	19.8	49.5	2066	4	US-09-248-796A-6306
23	19.8	49.5	2971	4	US-09-710-279-3449
24	19.8	49.5	9366	4	US-09-949-016-17437
25	19.8	49.5	74097	4	US-09-949-016-11785
26	19.8	49.5	74097	4	US-09-949-016-16239
27	19.8	49.5	265038	4	US-09-949-016-15779

28	19.6	49.0	189	4	US-09-540-236-1562	Sequence 1562, Ap
29	19.6	49.0	601	4	US-09-949-016-83110	Sequence 83110, A
30	19.6	49.0	601	4	US-09-949-016-83111	Sequence 83111, A
31	19.6	49.0	4322	2	US-08-537-342-1	Sequence 1, Appl1
32	19.6	49.0	6330	4	US-09-949-016-15506	Sequence 15506, A
33	19.6	49.0	66986	4	US-09-596-002-29	Sequence 29, Appl1
34	19.6	49.0	75480	4	US-09-949-016-16090	Sequence 16090, A
35	19.6	49.0	80858	4	US-09-949-016-12859	Sequence 12859, A
36	19.6	49.0	80859	4	US-09-949-016-15715	Sequence 15715, A
37	19.6	49.0	89892	4	US-09-949-016-13667	Sequence 13667, A
38	19.6	49.0	107937	4	US-09-949-016-17192	Sequence 17192, A
39	19.6	49.0	140315	4	US-09-949-016-14141	Sequence 14141, A
40	19.4	48.5	601	4	US-09-949-016-26227	Sequence 26227, A
41	19.4	48.5	601	4	US-09-949-016-74129	Sequence 74129, A
42	19.4	48.5	601	4	US-09-949-016-74160	Sequence 74160, A
43	19.4	48.5	601	4	US-09-949-016-202761	Sequence 202761, A
44	19.4	48.5	2810	4	US-09-949-016-5447	Sequence 5447, Ap
45	19.4	48.5	2816	3	US-09-643-597-333	Sequence 333, App

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Hultman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/880,499

FILING DATE: CONCURRENTLY HEREWITH

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match

Best Local Similarity 100.0%; Score 40; DB 3; Length 1394;

Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 40
|||
Db 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 1278

RESULT 2

US-08-880-499-2
Sequence 2, Application US/0880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garneat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 40; DB 3; Length 1394;

Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 40
|||
Db 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 1278

RESULT 3

US-09-248-796A-5546
Sequence 5546, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196,132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 5546
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-5546

Query Match 51.5%; Score 20.6; DB 4; Length 1071;

Best Local Similarity 74.3%; Pred. No. 51;
Matches 26; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 38
|||
Db 224 ATCCCACTCTCAACATGTCACATCTTACTTATCA 258

RESULT 4

US-09-949-016-14995/C
Sequence 14995, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14995
LENGTH: 42075
TYPE: DNA
ORGANISM: Human
US-09-949-016-14995

Query Match 51.5%; Score 20.6; DB 4; Length 42075;

Best Local Similarity 85.2%; Pred. No. 1.2e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTCA 30
|||
Db 31591 ATTCCACCAACAACATCATCTTCA 31565

RESULT 5

US-09-949-016-65770/C
Sequence 65770, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 65770
LENGTH: 601


```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-65770

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 601;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATCTACTCCCAACATCTTACTCATGCAA 39
Db 414 GGATCTCTGGCTCAGCATCTCTGCTCAGGCAA 377

RESULT 6
US-09-774-528-184
; Sequence 184, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyun
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619e1 Nucleic Acids and
; FILE REFERENCE: 802
; TITLE OF INVENTION: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pf_fl_genes Version 2.0
; SEQ ID NO 184
; LENGTH: 973
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (231)..(749)
US-09-774-528-184

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 973;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATCTACTCCCAACATCTTACTCATGCAA 39
Db 626 GGAAATCACTCCAAAGAAACCTCGAAACCATGCA 663

RESULT 7
US-08-945-056-4
; Sequence 4, Application US/08945056
; Patent No. 6077994
; GENERAL INFORMATION:
; APPLICANT: Coupland, George M.
; APPLICANT: Puterill, Joanna J.
; TITLE OF INVENTION: Genetic control of flowering
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 6077994th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,056
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02561
; FILING DATE: 01-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9422083.7
; FILING DATE: 02-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ms Mary J Wilson
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4201 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; STRAIN: Landsberg erecta
; POSITION IN GENOME:
; MAP POSITION: chromosome 5
US-08-945-056-4

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 3; Length 4201;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 GATACCTACTCCCAACATCTTACT 32
Db 2631 GATACGAGCTCCACACCATCAACTTACT 2660

RESULT 8
US-09-949-016-13668/c
; Sequence 13668, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13668
; LENGTH: 51770
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(51770)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13668

Query Match
51.0%; Score 20.4; DB 4; Length 51770;
```


Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 9851 GGATTCCTGGGCTCAAGCAATCCTCCTGCTCAGGCAA 9814

RESULT 9

US-09-949-016-17420

/ Sequence 17420, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

Query Match 51.0%; Score 20.4; DB 4; Length 84571;
Best Local Similarity 71.1%; Pred. No. 1.7e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 24559 GGAAATCAACTCCAAAGAAACCTTCAAAACCATGCAA 24596

RESULT 10

US-09-949-016-11824/C

/ Sequence 11824, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR FILING DATE: 2000-10-20

US-09-949-016-11824

Query Match 51.0%; Score 20.4; DB 4; Length 126200;

Best Local Similarity 71.1%; Pred. No. 1.9e+02;

Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 53133 GGAAATCAACTCCAAAGAAACCTTCAAAATCATGCAA 53096

RESULT 11

US-09-949-016-13193/C

/ Sequence 13193, Application US/09949016

/ Patent No. 6812339

/ GENERAL INFORMATION:

/ APPLICANT: VENTER, J. Craig et al.

/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

/ FILE REFERENCE: CL001307

/ CURRENT APPLICATION NUMBER: US/09/949,016

/ CURRENT FILING DATE: 2000-04-14

/ PRIOR APPLICATION NUMBER: 60/241,755

/ PRIOR FILING DATE: 2000-10-20

/ PRIOR APPLICATION NUMBER: 60/237,768

/ PRIOR FILING DATE: 2000-10-03

/ PRIOR APPLICATION NUMBER: 60/231,498

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

/ PRIOR FILING DATE: 2000-09-08

US-09-949-016-13193

Query Match 51.0%; Score 20.4; DB 4; Length 126200;

Best Local Similarity 71.1%; Pred. No. 1.9e+02;

Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 2 GGATACCTACTCCCAACAATCCATCTTACTCATGCAA 39

Db 53133 GGAAATCAACTCCAAAGAAACCTTCAAAATCATGCAA 53096

RESULT 12

US-09-610-040-5

/ Sequence 5, Application US/09610040

/ Patent No. 6465217

/ GENERAL INFORMATION:

/ APPLICANT: Boyes, Douglas

/ APPLICANT: Davis, Keith

/ APPLICANT: Moessner, Jeffrey

/ APPLICANT: Gorlach, Jorn

/ APPLICANT: Hamilton, Carol

/ APPLICANT: Hoffmann, Neil

/ APPLICANT: Kloci, Andreas

/ APPLICANT: Zayed, Adel

/ APPLICANT: Ascenzi, Robert

/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

/ IN PLANTS

/ FILE REFERENCE: 9128.14

/ CURRENT APPLICATION NUMBER: US/09/610,040

/ CURRENT FILING DATE: 2000-07-05

/ NUMBER OF SEQ ID NOS: 11

/ SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 5

/ LENGTH: 5099

/ TYPE: DNA


```
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4049)..(4049)
/ OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).
US-09-610-040-5

Query Match          50.5%; Score 20.2; DB 3; Length 5099;
Best Local Similarity 88.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      6 ACCTACTCCCAACATCCATCTTA 30
DB      2855 ACCTACCCCAACATCCCTTTA 2879

RESULT 13
US-10-267-763-5
/ Sequence 5, Application US/10267763
/ Patent No. 6800459
/ GENERAL INFORMATION:
/ APPLICANT: Goriach, Jörn
/ APPLICANT: Boyes, Douglas
/ APPLICANT: Davis, Keith
/ APPLICANT: Hamilton, Carol
/ APPLICANT: Hoffman, Neil
/ APPLICANT: Klotz, Andreas
/ APPLICANT: Woessner, Jeffrey
/ APPLICANT: Zayed, Adel
/ APPLICANT: Ascenzi, Robert
/ TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
/ TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
/ TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
/ FILE REFERENCE: 2035DIV1
/ CURRENT APPLICATION NUMBER: US/10/267,763
/ CURRENT FILING DATE: 2002-10-09
/ PRIOR APPLICATION NUMBER: 09/610,040
/ PRIOR FILING DATE: 2000-07-05
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 5
/ LENGTH: 5099
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(5099)
/ OTHER INFORMATION: n = A,T,C or G
US-10-267-763-5

Query Match          50.5%; Score 20.2; DB 4; Length 5099;
Best Local Similarity 88.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      6 ACCTACTCCCAACATCCATCTTA 30
DB      2855 ACCTACCCCAACATCCCTTTA 2879

RESULT 14
US-09-949-016-88655
/ Sequence 88655, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
```

```
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 88655
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-88655

Query Match          50.0%; Score 20; DB 4; Length 601;
Best Local Similarity 68.4%; Pred. No. 79;
Matches 26; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY      2 GGATACCTACTCCCAACATCCATCTTATCATGCA 39
DB      293 GTATAGTCTCCCAACGCTGCTTATATATAAA 330

RESULT 15
US-09-620-312D-22
/ Sequence 22, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yunging
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhimei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 22
/ LENGTH: 2615
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (160)..(849)
US-09-620-312D-22

Query Match          50.0%; Score 20; DB 4; Length 2615;
Best Local Similarity 72.2%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      4 ATACTACTCCCAACATCCATCTTACTCATGCA 39
DB      1855 ATACAGTTCCCAACCAACATCTTCATTGGAA 1890

Search completed: September 15, 2005, 08:25:23
Job time : 7.05048 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 07:10:11 ; Search time 26.0505 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatacctactcccccaacatccatcttaccatgcacac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	40	US-10-713-381-6	Sequence 6, Appli
2	40	100.0	50	US-10-713-381-5	Sequence 5, Appli
3	40	100.0	158	US-10-713-381-3	Sequence 3, Appli
4	40	100.0	255	US-10-713-381-9	Sequence 9, Appli
5	40	100.0	1394	US-10-713-381-1	Sequence 1, Appli
6	40	100.0	1394	US-10-713-381-2	Sequence 2, Appli
7	24	60.0	85779	US-10-713-993-6880	Sequence 6880, Ap

8	23.2	58.0	706	18	US-10-424-599-31006	Sequence 31006, A
9	22.6	56.5	294	20	US-10-425-115-143168	Sequence 143168, A
10	22.4	56.0	315	18	US-10-424-599-83179	Sequence 83179, A
11	22.4	56.0	226215	13	US-10-087-192-1948	Sequence 1948, A
12	21.6	54.0	4110	20	US-10-473-126-342	Sequence 342, App
13	21.6	54.0	6309	15	US-10-311-455-277	Sequence 277, App
14	21.4	53.5	1341	9	US-09-938-842A-1555	Sequence 1555, Ap
15	21.4	53.5	1341	11	US-09-938-842A-1555	Sequence 1555, Ap
16	21.4	53.5	6881	15	US-10-311-455-128	Sequence 128, App
17	21.4	53.5	6881	16	US-10-240-452-4	Sequence 4, Appli
18	21.4	53.5	8168	15	US-10-311-455-901	Sequence 901, App
19	21.4	53.5	8168	16	US-10-240-454-23	Sequence 23, Appli
20	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
21	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, Appli
22	21	52.5	293	18	US-10-424-599-24043	Sequence 24043, A
23	21	52.5	931	18	US-10-424-599-20320	Sequence 20320, A
24	21	52.5	2000	17	US-10-260-218-2594	Sequence 2594, App
25	21	52.5	2792	18	US-10-041-018-361	Sequence 361, App
26	21	52.5	3117	18	US-10-041-018-362	Sequence 362, App
27	21	52.5	5507	20	US-10-473-126-198	Sequence 198, App
28	21	52.5	5507	20	US-10-473-126-198	Sequence 198, App
29	21	52.5	8404	15	US-10-311-455-1568	Sequence 1568, App
30	21	52.5	8404	18	US-10-221-714A-222	Sequence 222, App
31	21	52.5	10286	14	US-10-239-676-13	Sequence 13, Appli
32	21	52.5	10286	15	US-10-240-453-21	Sequence 21, Appli
33	21	52.5	63294	11	US-09-997-722-205	Sequence 205, App
34	21	52.5	106315	13	US-10-087-192-178	Sequence 178, A
35	21	52.5	2140405	13	US-10-027-632-76212	Sequence 76212, A
36	21	52.5	2140405	17	US-10-027-632-76212	Sequence 76212, A
37	20.8	52.0	3263	19	US-10-437-963-4616	Sequence 4616, App
38	20.8	52.0	3292	13	US-10-027-632-112652	Sequence 112652, A
39	20.8	52.0	3292	17	US-10-027-632-112652	Sequence 112652, A
40	20.8	52.0	6620	14	US-10-239-676-196	Sequence 196, App
41	20.8	52.0	6620	15	US-10-240-453-290	Sequence 290, App
42	20.8	52.0	513309	10	US-09-754-853A-4	Sequence 4, Appli
43	20.6	51.5	970	18	US-10-424-599-91073	Sequence 91073, A
44	20.6	51.5	1103	22	US-10-424-599-91073	Sequence 91073, A
45	20.6	51.5	1889	22	US-10-499-352A-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6
Query Match 100.0%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCTCCCAACATCCATCTTACTCATGCAC 40

Db 1 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 40

RESULT 2
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match 100.0%; Score 40; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 40
Db 11 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 50

RESULT 3
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 100.0%; Score 40; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 40
Db 86 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 125

RESULT 4

US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 40; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 40
Db 99 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 138

RESULT 5
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 40
Db 1239 AGGATACCTACTCCCAAGCATCATCTTACTCATGCAAC 1278

RESULT 6
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.

APPLICANT: FOX, TIMOTHY W.
APPLICANT: GARNAT, CARL W.
APPLICANT: HUFMAN, GARY
APPLICANT: KENDALL, TIMMY L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
FILE REFERENCE: 578R
CURRENT APPLICATION NUMBER: US/10/713,381
CURRENT FILING DATE: 2003-11-14
PRIOR APPLICATION NUMBER: 08/880,499
PRIOR FILING DATE: 1997-06-23
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1394
TYPE: DNA
ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6,4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 1278

RESULT 7
US-10-719-993-6880/C
Sequence 6880, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: C1001496
CURRENT APPLICATION NUMBER: US/10/719,993
CURRENT FILING DATE: 2003-11-24
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6880
LENGTH: 85779
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(85779)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6880

Query Match 60.0%; Score 24; DB 20; Length 85779;
Best Local Similarity 84.4%; Pred. No. 38;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 4 ATACCTACTCCCAACATCATCTTACTCAT 35
DB 16082 AAACCTACTCCCAACATCATCTTACTCAT 16051

RESULT 8
US-10-424-599-31006
Sequence 31006, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 31006
LENGTH: 706
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_128002C.1
US-10-424-599-31006

Query Match 58.0%; Score 23.2; DB 18; Length 706;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 5 TACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 455 TCCCTTCTCCCAACCCCTCATCAACCTCATACAC 490

RESULT 9
US-10-425-115-143168
Sequence 143168, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 143168
LENGTH: 294
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_62049C.1
US-10-425-115-143168

Query Match 56.5%; Score 22.6; DB 20; Length 294;
Best Local Similarity 86.2%; Pred. No. 45;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 4 ATACCTACTCCCAACATCATCTTACT 32
DB 226 ATACCTGTCCCATACCTCATCTTACT 254

RESULT 10
US-10-424-599-83179
Sequence 83179, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 83179
LENGTH: 315
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179

Query Match 56.0%; Score 22.4; DB 18; Length 315;
Best Local Similarity 81.2%; Pred. No. 55;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 35
DB 211 AACGAACTCCTAAACATCATCTTAAATCCT 242

RESULT 11
US-10-087-192-1948
; Sequence 1948, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1948
; LENGTH: 226215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1948

Query Match 56.0%; Score 22.4; DB 13; Length 226215;
Best Local Similarity 72.5%; Pred. No. 2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 183206 AGAATTCCTGCTTCCAAACAGACATCTTACTCCTTTTAC 183245

RESULT 12
US-10-473-126-342/C
; Sequence 342, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE: proliferative disorders
; CURRENT APPLICATION NUMBER: US/10/473,126
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 342
; LENGTH: 4110
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-342

Query Match 54.0%; Score 21.6; DB 20; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 31
DB 3083 AACCTTAACCCCAACATCATCTTACTCAT 3056

RESULT 13

US-10-311-455-277/C
; Sequence 277, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIERENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 277
; LENGTH: 6309
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-277

Query Match 54.0%; Score 21.6; DB 15; Length 6309;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 39
DB 3335 ATACTATTCCCATACGACATCTCATATATCCAA 3300

RESULT 14
US-09-938-842A-1555/C
; Sequence 1555, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

Query Match 53.5%; Score 21.4; DB 9; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCA 39
DB 288 AGTCACCTATTCCCAACATCTCTGTGCTTACACAA 250

RESULT 15

US-09-938-842A-1555/c
 ; Sequence 1555, Application US/09938842A
 ; Publication No. US2004009476A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1555
 ; LENGTH: 1341
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1555

Query Match 53.5%; Score 21.4; DB 11; Length 1341;
 Best Local Similarity 71.8%; Pred. No. 1.9e+02;
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCA 39
 DB 288 AGTCCACTATTCCTCAACATCTGTGCTCACACAA 250

Search completed: September 15, 2005, 20:46:11
 Job time : 32.0505 sec

This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 124.754 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggataccctaccctcccaacacatccatctctcctcgaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	40	100.0	687	9	CC656939 OGM2020TV
2	40	100.0	915	9	CG224225 OGIAG08TV
3	40	100.0	963	9	CC656933 OGM2020TM
4	23.8	59.5	528	7	H24137 YMS0908.r1
5	23.6	59.0	375	8	B2706605 SM416B1-G
6	23.2	58.0	738	9	CO117589 CR181951
7	23.2	58.0	783	9	CR181951 Reverse 8
8	23.2	58.0	1018	9	CL109241 ISB1-51D1
9	23.2	58.0	1253	6	CD495749 CDA18-D05
C 10	23	57.5	405	8	AZ883408 RPCI-23-1
C 11	23	57.5	798	4	BG506830 601861443
C 12	23	57.5	968	4	BF143726 601789606
13	22.8	57.0	771	4	BM406466 EST580781
14	22.6	56.5	410	2	AW257178 EST305315
15	22.6	56.5	560	2	BE205271 EST397947
C 16	22.6	56.5	624	9	CR843833 GROAA78D
C 17	22.6	56.5	630	8	B2173201 CH230-369
C 18	22.6	56.5	644	8	CR843905 GROAA79A
C 19	22.6	56.5	664	8	AZ400686 IM0167P01
20	22.6	56.5	691	2	AW687128 NF006C08R
C 21	22.6	56.5	736	9	CR818130 GROAA43D
C 22	22.6	56.0	1080	9	CL058859 CH216-88P
23	22.4	56.0	298	2	BB264116 BB264116
C 24	22.4	56.0	940	7	CK864514 AGENCOURT

C 25	22.4	56.0	940	7	CK871857 AGENCOURT
C 26	22.4	56.0	951	7	CN317130 AGENCOURT
27	22.2	55.5	273	1	AI206216 q27f06.x
28	22.2	55.5	346	1	AI698201 wa67911.x
29	22.2	55.5	398	1	AA725064 a106h07.s
30	22.2	55.5	417	1	AI467879 l178e09.x
31	22.2	55.5	479	2	BF601210 266128 MA
32	22.2	55.5	491	1	AI382034 C633905.X
C 33	22.2	55.5	609	4	BM027586 G1T000109
C 34	22.2	55.5	727	7	CK949745 4074916 B
C 35	22.2	55.5	802	9	CG811817 FSAAM35TR
C 36	22.2	55.5	940	9	CG383910 OGA2P61TV
C 37	22.2	55.5	1090	9	CL084441 ISB1-3G16
C 38	22	55.0	207	7	CK913825 p3fmgcf-0
C 39	22	55.0	330	3	CNS098G7 BX071651 Single re
C 40	22	55.0	467	9	CG982704 CH240_164
C 41	22	55.0	478	3	CNS098GU BX047050 Single re
C 42	22	55.0	506	8	BX036328 Single re
C 43	22	55.0	506	8	AZ021739 RPCI-23-3
C 44	22	55.0	548	3	CNS09071 BX036328 Single re
C 45	22	55.0	576	8	AQ321478 RPCI11-10

ALIGNMENTS

RESULT 1	CC656939	687 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGM2020TV_ZM_0.7_1.5_KB	Zea mays	genomic clone	ZMMA0554D15,	
DEFINITION	genomic survey sequence.				
ACCESSION	CC656939				
VERSION	CC656939.1	GI:32060231			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: OGM2020TM Contact: Cathy Whitelaw TIGR				
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TF Class: sheared ends.				
source	Location/Qualifiers 1..687 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMA0554D15" /clone_11b="ZM_0.7_1.5_KB" /note="Vector: pBSGK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN	Query Match 100.0%; Score 40; DB 9; Length 687; Best Local Similarity 100.0%; Pred. No. 3.2e-05; Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40				
Db	337 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 298				

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS OGIAG08TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0716B15,
DEFINITION genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 915)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGIAG08TH
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0716B15"
/clone_1lb="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATCCATCTTACTATGCAAC 40
|||||
618 AGGATACCTACTCTCCAAACATCCATCTTACTATGCAAC 657

RESULT 3
CC656933 963 bp DNA linear GSS 19-JUN-2003
LOCUS OGMDQ20TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0554D15,
DEFINITION genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 963)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGMDQ20TV
Contact: Cathy Whiteaw
TIGR

FEATURES
source

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0554D15"
/clone_1lb="ZM 0.7 1.5 KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCAAACATCCATCTTACTATGCAAC 40
|||||
836 AGGATACCTACTCTCCAAACATCCATCTTACTATGCAAC 875

RESULT 4
H24137 528 bp mRNA linear EST 06-JUL-1995
LOCUS ym50g08.r1 Soares infant brain INIB Homo sapiens cDNA clone
DEFINITION IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1 GI:892832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 528)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaskis,E., Waterson,R., Williamson,A., Wohldmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1161
High quality sequence stops: 349
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_1lb="Soares infant brain INIB"
/note="Organ: whole brain; Vector: Latmid B4; Site_1: Not

1; Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15',
 AACTGAGAAATTCGGCGCGGCAATTTTCTTTTCTTTT 3');
 double-stranded cDNA was ligated to Hind III adaptors
 (Pharmacia), digested with Not I and directionally cloned
 into the Not I and Hind III sites of the Lactamid BA vector.
 Library went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 59.0%; Score 23.8; DB 7; Length 528;
 Best Local Similarity 80.0%; Pred. No. 97;
 Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 38
 |||||
 DB 346 AATCTACTCTCAACATTCATCTTATACATCA 380

RESULT 5

BZ706605/c 375 bp DNA linear GSS 18-MAY-2003
 LOCUS BZ706605
 DEFINITION SM416B1-G07_55_13.ab1 Spider Monkey genomic BAC library Atelles
 geoffroyi genomic, genomic survey sequence.

ACCESSION BZ706605
 VERSION BZ706605.1 GI:30843142

SOURCE Atelles geoffroyi (black-handed spider monkey)
 ORGANISM Atelles geoffroyi

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae;
 Atelles.

REFERENCE 1 (bases 1 to 375)
 Qian, Y., Jin, L. and Su, B.
 BAC end sequences of spider monkey genomic library
 Unpublished (2003)
 CONTACT: Qian Y., Jin L., Su B.
 CENTER for Genome Information
 University of Cincinnati
 Kettering Lab, 3223 Eden Ave., Cincinnati, OH 45267-0056, USA
 Tel: 1-513-558-6678
 Fax: 1-513-558-4505
 Email: subnucmail.uc.edu
 Seq primer: T7
 Class: BAC ends.

FEATURES

Location/Qualifiers
 1..375
 /organism="Atelles geoffroyi"
 /mol_type="genomic DNA"
 /strain="black-handed spider monkey"
 /db_xref="taxon:9509"
 /sex="Male"
 /issue_type="Skin"
 /cell_type="Fibroblast"
 /cell_line="AG05352"
 /dev_stage="3 DA"
 /clone_lib="Spider Monkey genomic BAC library"
 /note="Vector: pBACe3.6; Site 1: EcoRI; Genomic DNA was
 partially digested with EcoRI. Vector, pBACe3.6,
 Recombinants were transformed into DH10B."

ORIGIN
 source

Query Match 59.0%; Score 23.6; DB 8; Length 375;
 Best Local Similarity 74.4%; Pred. No. 1.1e+02;
 Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCATCTTACTCATGCA 40
 |||||
 DB 255 GGATATCTACTTGTGAACCATCATCTTACTNAGCAAC 217

RESULT 6

COL17589

LOCUS COL17589 738 bp mRNA linear EST 16-JUN-2004
 DEFINITION GR_EB01H15.r GR_Eb Gossypium raimondii cDNA clone GR_EB01H15 3',
 mRNA sequence.

ACCESSION COL17589
 VERSION COL17589.1 GI:48816276

KEYWORDS EST.
 SOURCE Gossypium raimondii

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE

1 (bases 1 to 738)
 Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C.,
 Udall, V.A., Rapp, R.A., Wendel, J.F., Rao, K., Soderlund, C. and
 Wang, R.A.

Global assembly of Cotton ESTs
 Unpublished (2004)
 CONTACT: Rod A. Wang
 Arizona Genomics Institute
 The University of Arizona
 Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
 Tel: 520 626 9595
 Fax: 520 621 1259
 Email: http://genome.arizona.edu

COMMENT

Plate: 01 row: H column: 15.
 Location/Qualifiers

FEATURES

source

1..738
 /organism="Gossypium raimondii"
 /mol_type="mRNA"
 /db_xref="taxon:29730"
 /clone="GR_EB01H15"
 /issue_type="floral"
 /dev_stage="3 to +3 DPA"
 /lab_host="GR_EB"
 /clone_lib="DH10B"
 /note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
 EcoRV; Library made by Invitrogen with RNA supplied by
 Wendie Lab. Directional cloned into NotI-EV. Clonies
 plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 58.0%; Score 23.2; DB 7; Length 738;
 Best Local Similarity 77.8%; Pred. No. 1.8e+02;
 Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 39
 |||||
 DB 224 ATACACACTCCCATACACGCAAGTACATCATGCA 259

RESULT 7

CR181951 783 bp DNA linear GSS 06-JUL-2004
 LOCUS CR181951
 DEFINITION Reverse strand read from insert in 5'HPT insertion targeting and
 chromosome engineering clone MHPN561p21, genomic survey sequence.

ACCESSION CR181951
 VERSION CR181951.1 GI:49960800
 GSS: genome survey sequence; MICEP.
 SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 783)
 Adams, D.J., Biggs, P.J., Cox, A.V., Davies, R.M., van der Weyden, L.,
 Jonkers, J., Smith, J., Plumb, R.W., Taylor, R.G., Nishijima, I., Yu, Y.,
 Rogers, J. and Bradley, A.

REFERENCE

AUTHORS

TITLE

JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICEP
 Location/Qualifiers

FEATURES

1..783
 /organism="Mus musculus"
 /mol_type="genomic DNA"

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: gzhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 189 row: L column: 2
Seq primer: SP6
class: BAC ends.

FEATURES

source
1..405
Location/Qualifiers

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-189L2"
/sex="Female"
/lab_host="DH10B"
/clone_1lb="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 57.5%; Score 23; DB 8; Length 405;
Best Local Similarity 74.4%; Pred. No. 1.9e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTTCTCCCAACATCCATCTTATCATGCAA 39
DB 182 AGGTACCAATCCCAACATCTTCTTATCATGCAA 144

RESULT 11
BG506830/c 798 bp mRNA linear EST 27-MAR-2001
LOCUS 601861443F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071154 5',
DEFINITION mRNA sequence.
ACCESSION BG506830
VERSION BG506830.1 GI:13468347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 798)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM916 row: e column: 11
High quality sequence stop: 2.
Location/Qualifiers

FEATURES

source
1..798
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:4071154"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1lb="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccatcgcc); Site 2: SfiI (ggccatcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGACATATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 57.5%; Score 23; DB 4; Length 798;
Best Local Similarity 83.9%; Pred. No. 2.2e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 ACTCCCAACATCCATCTTACTCATGCAAC 40
DB 222 ACTCCCAACAGCCTTCTTATCATGCAATC 192

RESULT 12
BF143726 968 bp mRNA linear EST 24-OCT-2000
LOCUS 601789606F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4020600 5',
DEFINITION mRNA sequence.
ACCESSION BF143726
VERSION BF143726.1 GI:10982766
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 968)
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM9274 row: k column: 01
High quality sequence start: 3
High quality sequence stop: 648.
Location/Qualifiers

FEATURES

source
1..968
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:4020600"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Lu30"
/note="Organ: lung; Vector: PCMV-SPORE6; Site 1: NotI; Site 2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 57.5%; Score 23; DB 2; Length 968;
Best Local Similarity 74.4%; Pred. No. 2.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACAATCATCTTACTCATGCAAC 40
 |||||
 DB 818 GGCTCCGTACTCCCAACAACCAACCCACACACCCCAAC 856

RESULT 13
 BM406466
 LOCUS 771 bp mRNA linear EST 10-MAR-2003
 DEFINITION EST580781 potato roots Solanum tuberosum cDNA clone cPROZ7F14.5'
 end, mRNA sequence.
 ACCESSION BM406466
 VERSION BM406466.1 GI:18258084
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 771)
 van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Utechtback, T., Chiemingo, A., Bougri, O., Buel, C.R., Romning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buel
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3.

FEATURES
 source
 Location/Qualifiers
 1..771
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPROZ7F14"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab.
 sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN

Query Match 57.0%; Score 22.8; DB 4; Length 771;
 Best Local Similarity 79.4%; Pred. No. 2.6e+02;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACAATCATCTTACTCATGCA 38
 |||||
 DB 649 TCCCTCTCTCCACCAATCCACATCTCTCAAGCA 682

RESULT 14
 AM257178
 LOCUS 410 bp mRNA linear EST 20-DEC-1999
 DEFINITION EST305315 KV2 Medicago truncatula cDNA clone KV2-7D15, mRNA
 sequence.
 ACCESSION AM257178
 VERSION AM257178.1 GI:6605435
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 ESTs from roots of Medicago truncatula after Rhizobium inoculation
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenbosch.umn.edu
 Texas A&M EST name: T115883e
 TIGR sequence name: MTAAO207K
 More information is available at: . (and for clone ordering info)
 http://chryslr.tamu.edu/medicago
 Seq primer: SKmod (CTA GAA CTA gtc gAT CC).
 Location/Qualifiers
 1..410
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="KV2-7D15"
 /tissue_type="Seedling roots"
 /dev_stage="2 days post-inoculation with Sinorhizobium meliloti"
 /lab_host="E. coli strain SOLR"
 /clone_lib="KV2"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Uniap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Query Match 56.5%; Score 22.6; DB 2; Length 410;
 Best Local Similarity 86.2%; Pred. No. 2.8e+02;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ACTTACTCCCAACAATCATCTTACTCA 34
 |||||
 DB 16 ACCAAGCTCCCAACAATCATCTGCTCA 44

RESULT 15
 BE205271
 LOCUS 560 bp mRNA linear EST 05-SEP-2000
 DEFINITION EST397947 KV0 Medicago truncatula cDNA clone pKV0-21123, mRNA
 sequence.
 ACCESSION BE205271
 VERSION BE205271.1 GI:8748567
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 1 (bases 1 to 560)
 Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Frazer, C.M.
 ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738

Email: kvanden@cbs.umn.edu
 Texas A&M University name: T265043e
 TIGR sequence name: MTGBK60TK

More information is available at: .

<http://chryslr.tamu.edu/medicago>
 Seq primer: Skmod (CTA gaa cta gtc gat cc).

FEATURES

source

```

1..560
    Location/Qualifiers
        /organism="Medicago truncatula"
        /mol_type="mRNA"
        /cultivar="genotype A17"
        /db_xref="taxon:3880"
        /clone="pKV0-21123"
        /issue_type="Seedling roots"
        /dev_stage="Immediately prior to inoculation with
        Sinorhizobium meliloti (0 hour)"
        /lab_host="E.coli strain XL0LR"
        /clone_lib="KV0"
        /note="Vector: Bluescript SK-; Site 1: EcoRI; Site 2:
        XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
        was directionally ligated into the Unizap XR vector from
        Stratagene and packaged using Gigapack III Gold packaging
        extracts. Plasmids containing cDNA inserts were excised
        from the recombinant Lambda-Zap phage using Ex-assist
        helper phage and propagated in XL0LR cells."

```

ORIGIN

Query Match

Best Local Similarity 56.5%; Score 22.6; DB 2; Length 560;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ACCTACTCCCAACATCCATCTTACTCA 34
 |||||
 DB 24 ACCACTCCCAACATCCATCTTACTCA 52

Search completed: September 16, 2005, 08:08:43
 Job time : 124.754 sec

This page blank (uspio)

Db 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCCATCAAGGCGCTTCGATGGCCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTGTATCCAGGGGATATATGCCCCACAAATGCTACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGTATCCAGGGGATATATGCCCCACAAATGCTACCTATA 300
Qy 301 TTATATTTCTTTAGATATTTATTTTGAAGAAATAAACAATTATATCTTTGTGTA 360
Db 301 TTATATTTCTTTAGATATTTATTTTGAAGAAATAAACAATTATATCTTTGTGTA 360
Qy 361 GGGGCTTCAGATAGATTTTCGCTTAAAGGCCAGAAATCGAGGACCAAGCCATCTTAAGT 420
Db 361 GGGGCTTCAGATAGATTTTCGCTTAAAGGCCAGAAATCGAGGACCAAGCCATCTTAAGT 420
Qy 421 TCACATTTGGGACTCCGAGACAAAGATTAAATAATTAACAAGTAATCTATTCCTACT 480
Db 421 TCACATTTGGGACTCCGAGACAAAGATTAAATAATTAACAAGTAATCTATTCCTACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCACGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCACGATCCTCTTAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTCAAAACATCTTAAGCGACAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTCAAAACATCTTAAGCGACAA 600
Qy 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Qy 661 ATTGTTTTATATACATTTTCTTCTCAATAGAGTATTTCTTCGATTTTAA 720
Db 661 ATTGTTTTATATACATTTTCTTCTCAATAGAGTATTTCTTCGATTTTAA 720
Qy 721 ATGACTATTAAGCTATTTTATTAAGAGACGCAATGCTGATCTCTGTTCAAAAATC 780
Db 721 ATGACTATTAAGCTATTTTATTAAGAGACGCAATGCTGATCTCTGTTCAAAAATC 780
Qy 781 TTTCGATTTTAAAGCTATTTGGCAACCTGTTCTTCAAGAATTTGATTT 840
Db 781 TTTCGATTTTAAAGCTATTTGGCAACCTGTTCTTCAAGAATTTGATTT 840
Qy 841 TTCAAAAAAATTAATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATTAATTTTCTCTTATTAATAAGAAAACATTAGAAAAATAGAGT 900
Qy 901 TGGCAACTAGCCCTGAGATGTTTCCCAATTAATTAACAATCTGTGTAATTAATTTG 960
Db 901 TGGCAACTAGCCCTGAGATGTTTCCCAATTAATTAACAATCTGTGTAATTAATTTG 960
Qy 961 GCCAGCCCATTAATTTATTTAAACGAAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCATTAATTTATTTAAACGAAACTGAATCGAGCGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCACTTTTAAATGCAATGTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAAATCACTTTTAAATGCAATGTGCC 1080
Qy 1081 TGAGATGTGCGGTTTGGCAACATAGCCACCTGATATCATAGCTGAGTGTGCTAGCTA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACATAGCCACCTGATATCATAGCTGAGTGTGCTAGCTA 1140
Qy 1141 GGTTCGGAGCTCTCGTGCATCTCAATGAGCATATGCTTGTCAACCGTTCGTC 1200
Db 1141 GGTTCGGAGCTCTCGTGCATCTCAATGAGCATATGCTTGTCAACCGTTCGTC 1200
Qy 1201 TTGTTTCATCTGTCAGAGCCTTGGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
Db 1201 TTGTTTCATCTGTCAGAGCCTTGGCTATTTGAAACCAAGAGATACCTATCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGTAAC 1311
Db 1261 CCATCTTACTCATGCAACTTCCATGCAAAACGACATATGTTTCTGTAAC 1311

RESULT 2
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 100.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.6e-252; Indels 0; Gaps 0;
Matches 1311; Conservative 0; Mismatches 0;
Qy 1 CCATGTCCTCTATGAAAAAGATAGTACATGTCATATCCGTTTCTTAAAGGTC 60
Db 1 CCATGTCCTCTATGAAAAAGATAGTACATGTCATATCCGTTTCTTAAAGGTC 60
Qy 61 CTTCCTTGCCTTATTAATCTGATGATCGGGTTACAAAAAATTCCACGCGTGCATAT 120
Db 61 CTTCCTTGCCTTATTAATCTGATGATCGGGTTACAAAAAATTCCACGCGTGCATAT 120
Qy 121 CTCGATGTTTCACTCTCCACCTCGGCTGACATTTCTTGGAGTGTGAGTCCCAT 180
Db 121 CTCGATGTTTCACTCTCCACCTCGGCTGACATTTCTTGGAGTGTGAGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCGCTTTCGATGAGCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCGCTTTCGATGAGCCA 240
Qy 241 CGAGAGCTATGCGGTCGTGTGATCCAGGGATATATGTCCTCCCAATTCGTACCTATA 300
Db 241 CGAGAGCTATGCGGTCGTGTGATCCAGGGATATATGTCCTCCCAATTCGTACCTATA 300
Qy 301 TTATATTTCTTTAGATATTTATTTTGAAGAAATTAACAATCTTATCTTTGTGTA 360
Db 301 TTATATTTCTTTAGATATTTATTTTGAAGAAATTAACAATCTTATCTTTGTGTA 360
Qy 361 GGGGCTTCAGATAGATTTTTCGTTAGGGCCAGAAATCGAGGACCAAGCTGTCTAGTG 420
Db 361 GGGGCTTCAGATAGATTTTTCGTTAGGGCCAGAAATCGAGGACCAAGCTGTCTAGTG 420
Qy 421 TCACATTTGGGACTACCCAGACAAAGATTTAAAAAATAACAAGTAATCAATCCACT 480
Db 421 TCACATTTGGGACTACCCAGACAAAGATTTAAAAAATAACAAGTAATCAATCCACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCACGATCCTTAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCACGATCCTTAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTCAAAACATCTTAAGGCGACAAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATATGTTTACAGTTTCAAAACATCTTAAGGCGACAAA 600
Qy 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660
Db 601 TTATATCGAAGGTAAGCTATGAGTTCAGATTTTCTTTTCAATCTGTTATTTGTT 660

QY 661 ATTGTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATAA 720
 DB 661 ATTTGTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATAA 720
 QY 721 ATGACTATAAGTCATTTTATATAGAGACGCGATGCTGATGATTTCTCGTCAAAAATC 780
 DB 721 ATGACTATAAGTCATTTTATATAGAGACGCGATGCTGATGATTTCTCGTCAAAAATC 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATATAAATGAAAACCTTAGAAAATAGAT 900
 DB 841 TTCAAAAAAATTAAGTTATTTCTCTTATATAAATGAAAACCTTAGAAAATAGAT 900
 QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGATTAATTTG 960
 DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGATTAATTTG 960
 QY 961 GCCAGCCCATTAATTTATTAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCATTAATTTATTAACCGAACTGAATGAGCGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTTAGATTTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 DB 1021 TTCTTAGATTTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
 QY 1081 TGAGATGTCGGGTTTGSCAAGATAGCCACCGTATGATGATGATGATGATGATGAT 1140
 DB 1081 TGAGATGTCGGGTTTGSCAAGATAGCCACCGTATGATGATGATGATGATGATGAT 1140
 QY 1141 GGTTCGGACCTCTCGTGTATCTGACATGCGATGATGATGATGATGATGATGAT 1200
 DB 1141 GGTTCGGACCTCTCGTGTATCTGACATGCGATGATGATGATGATGATGATGAT 1200
 QY 1201 TTGTTCCATGTCGCAAGCCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCTTGGCTATTTCTGAACCAAGAGATACCTATCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCACTTCATGCAACCGACATATGTTTCTGAAAC 1311
 DB 1261 CCATCTTACTCATGCACTTCATGCAACCGACATATGTTTCTGAAAC 1311
 RESULT 3
 BD062176 1394 bp DNA linear PART 27-AUG-2002
 LOCUS BD062176 Male tissue-preferred regulatory region and method of using same.
 DEFINITION BD062176.1 GI:2260781
 ACCESSION BD062176.1
 VERSION JP 2001520523-A/1
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1394)
 AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
 COMMENT PIONEER HI BRED INTERNATIONAL INC
 PN JP 2001520523-A/1
 PD 30-OCT-2001
 PF 19-JUN-1998 JP 199504910
 PR 23-JUN-1997 US 08/880499
 PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN,
 PI TIMMY L KENDALL
 PC C12N15/82, C12N15/29, C12N9/24, C12N9/10, C12N9/00 PC
 PC A01H5/00
 CC Strandedness: Single;

CC Topology: Linear;
 FH Key Location/Qualifiers.
 FEATURES
 source 1..1394
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 100.0%; Score 1311; DB 6; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 3, 6e-252;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGAGGTC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGAGGTC 60
 QY 61 CTTCCTGCTGCTTATGACTGATGAAATCGGGGTTACAAAAAATTCCAGGGTGATGAT 120
 DB 61 CTTCCTGCTGCTTATGACTGATGAAATCGGGGTTACAAAAAATTCCAGGGTGATGAT 120
 QY 121 CTCATGTTCCACTTCTCCACCTCGGTTGCACTTTCTTGAGTGTGGTGGTCCAT 180
 DB 121 CTCATGTTCCACTTCTCCACCTCGGTTGCACTTTCTTGAGTGTGGTGGTCCAT 180
 QY 181 CTGACCGAGGCCCATGACACACTTTGCGGACACCCATCAAGGCGCTTTCGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATGACACACTTTGCGGACACCCATCAAGGCGCTTTCGATGGCCCA 240
 QY 241 CGAGAGTATCGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 CGAGAGTATCGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 TTATTAATCTTATGATTAATTAATTTTGGAAAAATACAACTTAATCTTTGATG 360
 DB 301 TTATTAATCTTATGATTAATTAATTTTGGAAAAATACAACTTAATCTTTGATG 360
 QY 361 GGGGCTCAGATGATTTTTCGTTAGGCGCCAGAAATGCGAGACACAGCCATGCTAGTG 420
 DB 361 GGGGCTCAGATGATTTTTCGTTAGGCGCCAGAAATGCGAGACACAGCCATGCTAGTG 420
 QY 421 TCCACTATTTGGGACTCCAGAACAGATTTTAAAAAATTAACCAAGTAACTTAATCCACT 480
 DB 421 TCCACTATTTGGGACTCCAGAACAGATTTTAAAAAATTAACCAAGTAACTTAATCCACT 480
 QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGCATCTTTAAAAAA 540
 DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGCATCTTTAAAAAA 540
 QY 541 CAAGCATTTTGAAGAGACAAATTAATGTTTACAGTTTCAAAACATCTAAGGCGCAAA 600
 DB 541 CAAGCATTTTGAAGAGACAAATTAATGTTTACAGTTTCAAAACATCTAAGGCGCAAA 600
 QY 601 TTATATGAAAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 DB 601 TTATATGAAAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
 QY 661 ATTTGTTTATATACATTTTCTTCTTACAAATAGAGATTTTCTTCGATTTTAAAA 720
 DB 661 ATTTGTTTATATACATTTTCTTCTTACAAATAGAGATTTTCTTCGATTTTAAAA 720
 QY 721 ATGACTATAAGTCATTTTATATAGAGACGCGATGCTGATGATTTCTCGATTTTAAAA 780
 DB 721 ATGACTATAAGTCATTTTATATAGAGACGCGATGCTGATGATTTCTCGATTTTAAAA 780
 QY 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTATAGAGCTAGTTGGCAACCCGTGTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAATTAAGTTATTTCTCTTATATAAATGAAAACCTTAGAAAATAGAT 900
 DB 841 TTCAAAAAAATTAAGTTATTTCTCTTATATAAATGAAAACCTTAGAAAATAGAT 900

Qy	901	GGCAGACTACGCCCTAGAAATGTTTTCCAAATTAATTAATCACTGCTAATAATTTATTTG	960
Dp	901	TGCCAGACTACGCCCTAGAAATGTTTTCCAAATTAATTAATCACTGCTAATAATTTATTTG	960
Qy	961	GCCAGCCCCCATTAATTTATTTAAACGGAATCTGAATCGAGGAAACCAATCTGAGCTAT	1020
Dp	961	GCCAGCCCCCATTAATTTATTTAAACGGAATCTGAATCGAGGAAACCAATCTGAGCTAT	1020
Qy	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCAGTTTAACTCATTTGTCCTCC	1080
Dp	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAATCAGTTTAACTCATTTGTCCTCC	1080
Qy	1081	TGAATATGCGCGTTTGGCAAGATAGCCACCGTATCTACTACTATAGTGTGCTCACTGCA	1140
Dp	1081	TGAATATGCGCGTTTGGCAAGATAGCCACCGTATCTACTACTATAGTGTGCTCACTGCA	1140
Qy	1141	GGTTCCGAGAGCTCGTGTCACTCAATGCGATCTCACTGCTGTTCCAACCGTTGCTC	1200
Dp	1141	GGTTCCGAGAGCTCGTGTCACTCAATGCGATCTCACTGCTGTTCCAACCGTTGCTC	1200
Qy	1201	TTGTTCCATCGTCCAAAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCAAACAT	1260
Dp	1201	TTGTTCCATCGTCCAAAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCAAACAT	1260
Qy	1261	CCATCTTACTATGCAATCTTCATGCAAAACGCAATATGTTTCTCGAAC	1311
Dp	1261	CCATCTTACTATGCAATCTTCATGCAAAACGCAATATGTTTCTCGAAC	1311

RESULT 4				
BD062177				
LOCUS	BD062177	1394 bp	DNA	linear PAT 27-AUG-2002
DEFINITION	Male tissue-preferred regulatory region and method of using same.			
ACCESSION	BD062177			
VERSION	BD062177.1	GI:22607782		
KEYWORDS	JP 2001520523-A/2.			
SOURCE	Homo sapiens	(human)		
ORGANISM	Homo sapiens			
REFERENCE	Enkayota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 1394) Albertsen,M.C., Fox,T.W., Garnaat,C.W., Hufman,G.A. and Kendall,T.L.			
TITLE	Male tissue-preferred regulatory region and method of using same			
JOURNAL	Patent: JP 2001520523-A 2 30-OCT-2001;			
COMMENT	PIONEER HI BRED INTERNATIONAL INC PN JP 2001520523-A/2 PD 30-OCT-2001 PF 19-JUN-1998 JP 1998504910 PR 23-JUN-1997 US 08/880499 PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN, PI TIMMY L KENDALL PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC CO7K14/34,C12Q1/68, PC A01H5/00 CC Strandedness: Single; CC Topology: linear; FH Key Location/Qualifiers. location/Qualifiers 1..1394 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"			
FEATURES				
source				
ORIGIN				

```

Query Match      100.0%; Score 1311; DB 6; Length 1394;
Beet Local Similarity 100.0%; Pred. No. 3.6e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1  CCATGGTGTCTCTATGAAAAAGATGAGTACATGTGCTATATCCGTTTTCTTAGGGTCC 60
      |||
Db      1  CCAATGGTGTCTCTATGAAAAAGATGAGTACATGTGCTATATCCGTTTTCTTAGGGTCC 60

```

QY	61	CTTCTTCGCTCTATTA	CTGA	CTGAATGGGGGTAC	AAAAAACTTCCACGGGTGCATGAT	120	
Db	61	CTTCTTCGCTCTATTA	CTGA	CTGAATGGGGGTTC	AAAAAACTTCCACGGGTGCATGAT	120	
QY	121	CTCCATGTTCCACTCT	CTCCCACTCGCGGTG	CACATTTCTTGGATG	TGCGGTGCCAT	180	
Db	121	CTCCATGTTCCACTCT	CTCCCACTCGCGGTG	CACATTTCTTGGATG	TGCGGTGCCAT	180	
QY	181	CTGACCGAGGCCCAT	CAGACACCTTTG	GGGACACCCATC	MAAGGCCCTTTCGATG	240	
Db	181	CTGACCGAGGCCCAT	CAGACACCTTTG	GGGACACCCATC	MAAGGCCCTTTCGATG	240	
QY	241	CGAAGACGTA	TGGGGTGGTGAT	CTCAGGGGATATATG	TCCCCCACAATCGTCACTATA	300	
Db	241	CGAAGACGTA	TGGGGTGGTGAT	CTCAGGGGATATATG	TCCCCCACAATCGTCACTATA	300	
QY	301	TTAATAATCTTTAGAT	TTAATAATTTTGG	AAAAAATAACAA	CTTATACCTTTG	360	
Db	301	TTAATAATCTTTAGAT	TTAATAATTTTGG	AAAAAATAACAA	CTTATACCTTTG	360	
QY	361	GGGGCTCAGACATAG	ATTTTCGTTA	GGGCCCCGAAATG	CGAGACACGCTATAGT	420	
Db	361	GGGGCTCAGACATAG	ATTTTCGTTA	GGGCCCCGAAATG	CGAGACACGCTATAGT	420	
QY	421	TCCACTATATGG	CACTACCCAGAACAGATTT	AAAAAAATACCA	AGTAATCACTACACT	480	
Db	421	TCCACTATATGG	CACTACCCAGAACAGATTT	AAAAAAATACCA	AGTAATCACTACACT	480	
QY	481	CGAAAGCATATCATG	TAATGTTTAA	AGAAACATCTATTA	AAACACGATCTCTTAAAA	540	
Db	481	CGAAAGCATATCATG	TAATGTTTAA	AGAAACATCTATTA	AAACACGATCTCTTAAAA	540	
QY	541	CAAGCATATTT	CGAAAGACAAATATATG	TATACAGTTTACAA	CAACTTAAGACGCA	600	
Db	541	CAAGCATATTT	CGAAAGACAAATATATG	TATACAGTTTACAA	CAACTTAAGACGCA	600	
QY	601	TTATATCGAA	AGCTAAGCTATG	ACGTTTCTTTCATTTCT	TCTTGTATTTTGT	660	
Db	601	TTATATCGAA	AGCTAAGCTATG	ACGTTTCTTTCATTTCT	TCTTGTATTTTGT	660	
QY	661	ATTGTTTTATAT	CAATTTCTTCCTTAC	AAATAGAGATTTCT	CCGATTTTATA	720	
Db	661	ATTGTTTTATAT	CAATTTCTTCCTTAC	AAATAGAGATTTCT	CCGATTTTATA	720	
QY	721	ATGACTATAA	AGTCATTTTATAT	TATNAGACGCA	TGTCGTATTCGTTCA	780	
Db	721	ATGACTATAA	AGTCATTTTATAT	TATNAGACGCA	TGTCGTATTCGTTCA	780	
QY	781	TTTCTGATTTT	TATNAGCTAG	TTGGCAACCTGT	TTCTTCAAGAAATTTGATTT	840	
Db	781	TTTCTGATTTT	TATNAGCTAG	TTGGCAACCTGT	TTCTTCAAGAAATTTGATTT	840	
QY	841	TTTCAAAAAA	ATATGTTTATTTCT	TTTATNAAATATG	AAAAACCTTATG	900	
Db	841	TTTCAAAAAA	ATATGTTTATTTCT	TTTATNAAATATG	AAAAACCTTATG	900	
QY	901	TGCGACATAG	CCCTAGATGTTT	CCCAATAATTA	CAATCACTGTGATATATTTG	960	
Db	901	TGCGACATAG	CCCTAGATGTTT	CCCAATAATTA	CAATCACTGTGATATATTTG	960	
QY	961	GCCAGCCCCCAT	ATAATTTATA	CCGAACTG	AAATCGAGCAACCAATCTG	1020	
Db	961	GCCAGCCCCCAT	ATAATTTATA	CCGAACTG	AAATCGAGCAACCAATCTG	1020	
QY	1021	TTTCTCTAG	ATTAATTA	AGGAGAGAGAA	GAATCACTTTTAAGTCA	1080	
Db	1021	TTTCTCTAG	ATTAATTA	AGGAGAGAGAA	GAATCACTTTTAAGTCA	1080	
QY	1081	TGAATATG	CGGTATGG	CAACGATAG	CCAGCTATCA	TAGCTATAGGCTCTAC	1140
Db	1081	TGAATATG	CGGTATGG	CAACGATAG	CCAGCTATCA	TAGCTATAGGCTCTAC	1140
QY	1141	GGTTTCG	AGCTCTCG	TGCTCATCT	CACTGATCTA	CTAGCTTGTTC	1200

Db 1141 GGTTCGGACGCTCTCGTGTATCTCAATGCAATGCTTGTTCACCGCTTCGTC 1200
Qy 1201 TTGTTCATCGTCCAAAGCCTTGTCTATTCGAAACCAAGAGATACCTACCCAAACAT 1260
Db 1201 TTGTTCATCGTCCAAAGCCTTGTCTATTCGAAACCAAGAGATACCTACCCAAACAT 1260
Qy 1261 CCATCTTACTACGCAACTTCGATGCAACAGCAGCATATGTTTCTCGAAC 1311
Db 1261 CCATCTTACTACGCAACTTCGATGCAACAGCAGCATATGTTTCTCGAAC 1311
RESULT 5
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS Zea mays male fertility protein (M45) gene, complete cds.
DEFINITION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogonaceae; Zea.
REFERENCE 1 (bases 1 to 3343)
Fox, T.W., Trimell, M.R. and Albertsen, M.C.
TITLE Cloning of M45, a gene required for male fertility from Zea mays
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 3343)
Fox, T.W., Trimell, M.R. and Albertsen, M.C.
REFERENCE Direct Submission
AUTHORS Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
TITLE Hi-Bred Inc., Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
JOURNAL IA 50131-1004, USA
FEATURES
source Location/Qualifiers
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/cui="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..32942
/gene="M45"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="M45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="M45"
/codon_start=1
/product="male fertility protein"
/protein_id="AKS2489.1"
/db_xref="GI:14028757"
/translation="MEKRNLDNRGRDGIQVPHLFEALALALVADPFGSLAEV
DVRPVHBLAPGEVWGMWSRLRGRLFGVEVGEVPSIEEDLGRGVNAGLA
DGRVAVWMBEAEKETPAVYVNPMSSEVCANGVNSTRRQHEKEEBCGRPLGRIGE
TGLVYADAVYGLVWQSGVAVSARELDGPRIIPANDLVHNGSVFDTDSRY
SRDHLNILLBEGTGLRLYDEPISGVHVLKGLVFPNGVOISDHOGLLSEITNC
RIRWVLEGRAGEVFANLPGFPNVNSNGRGQVVAIDCRRTPAQGVFAPLR
TLVFKPLSLKVLTKWAKRMHTVLLALDDEGRVVEVLEDRGHEVVKLVSEVREVRK
LWIGTVAHHNIATIPYLED"

ORIGIN

Query Match 100.0%; Score 1311; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 3.5e-252;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 CTCATGTTCCACTTCTCCACCTCGCGTTGCAATTTCTTGAGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTCGCGTTGCAATTTCTTGAGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTGGGAGACCCATCAAGGACCTTTCGATGAGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTGGGAGACCCATCAAGGACCTTTCGATGAGCCCA 240
Qy 241 CGAGAGTATCGGGTGTGTGATATCCAGGGATATATGTCCCCCAATGTCACCTATA 300
Db 241 CGAGAGTATCGGGTGTGTGATATCCAGGGATATATGTCCCCCAATGTCACCTATA 300
Qy 301 TTATTTATCTTATGATATTTATTTATTTTGGAAAAATACAACTTAATCTTTGTGTA 360
Db 301 TTATTTATCTTATGATATTTATTTATTTTGGAAAAATACAACTTAATCTTTGTGTA 360
Qy 361 GGGGCTCAGCATATGATTTTGGCTTAGGGCCGAGAAATGCGAGGACCAAGCATGTCTAGTG 420
Db 361 GGGGCTCAGCATATGATTTTGGCTTAGGGCCGAGAAATGCGAGGACCAAGCATGTCTAGTG 420
Qy 421 TCACATATTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAATCAATCACT 480
Db 421 TCACATATTGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAATCAATCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGCATCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAAACCAAGCATCTCTTAAAAA 540
Qy 541 CAAGCATATTTGAAAGAGACAAATATGTTACAGTTTCAAAACATTAAGGCGACAA 600
Db 541 CAAGCATATTTGAAAGAGACAAATATGTTACAGTTTCAAAACATTAAGGCGACAA 600
Qy 601 TTATATCGAAGAGTAAGCTATGACGTTTCAGATTTTCTTTTCATCTTGTATTTGTT 660
Db 601 TTATATCGAAGAGTAAGCTATGACGTTTCAGATTTTCTTTTCATCTTGTATTTGTT 660
Qy 661 ATTTGTTTATATATCAATTTTCTTCTTCAATAGATATTTCTTCCGATTTTATAA 720
Db 661 ATTTGTTTATATATCAATTTTCTTCTTCAATAGATATTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATAAGCATTTTATATTAAGACAGCATGTCGTGATTCGTGCAAAATC 780
Db 721 ATGACTATAAGCATTTTATATTAAGACAGCATGTCGTGATTCGTGCAAAATC 780
Qy 781 TTTCTGATTTTAAAGACTAGTTTGGCAACCTGTCTTTTCAAAAGATTTGATTT 840
Db 781 TTTCTGATTTTAAAGACTAGTTTGGCAACCTGTCTTTTCAAAAGATTTGATTT 840
Qy 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAATGAACACTTAGAAAAATAGAGT 900
Db 841 TTCAAAAAAATTAAGTTATTTTCTTATTAATAATGAACACTTAGAAAAATAGAGT 900
Qy 901 TCCGAGCATAGCCCTAGAAATGTTTCCCAATAATTAACAATCAGTGTATTAATTTTG 960
Db 901 TCCGAGCATAGCCCTAGAAATGTTTCCCAATAATTAACAATCAGTGTATTAATTTTG 960
Qy 961 GCGAGCCCATTAATTTAAACGAACTGAAATCGAGCAAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTTAAACGAACTGAAATCGAGCAAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAAGTAATAAGGAGAGAGAGAGAAATCAATTTAAGTCAATGTCCTC 1080
Db 1021 TTCTCTAGATTAAGTAATAAGGAGAGAGAGAGAAATCAATTTAAGTCAATGTCCTC 1080
Qy 1081 TGAGATGTCGGGTTGGCAAGATAGCCACGTAATCATAGGCTATAGGCTCTAGTCA 1140
Db 1081 TGAGATGTCGGGTTGGCAAGATAGCCACGTAATCATAGGCTATAGGCTCTAGTCA 1140
Qy 1141 GGTTCGGACGCTCTCGTGTATCTCAATGAGATCAATGCTTTGCAACCGTTTCGTC 1200
Db 1141 GGTTCGGACGCTCTCGTGTATCTCAATGAGATCAATGCTTTGCAACCGTTTCGTC 1200

QY 1201 TTGTCATGCTGCAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAAT 1260
|||||
Db 1201 TTGTCATGCTGCAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAAT 1260
|||||
QY 1261 CCATCTTACTGATGCAACTTTCATGCAACGCAATATGTTCTCTGAAC 1311
|||||
Db 1261 CCATCTTACTGATGCAACTTTCATGCAACGCAATATGTTCTCTGAAC 1311
|||||
RESULT 6
AX224402 AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 9 from Patent WO0160997.
ACCESSION AX224402
VERSION AX224402.1 GI:15554644
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
LOCATION/Qualifiers
1. 1255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 12.3%; Score 160.6; DB 6; Length 255;
Best Local Similarity 97.6%; Pred. No. 5.9e-22;
Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1145 CGGACACTCTGCTGATCTCATGATGCAATGCTGTTGCAACGGTTGCTGTTG 1204
|||||
Db 5 CCGCGGATCCCGGTGATCTCATGATGCAATGCTGTTGCAACGGTTGCTGTTG 64
|||||
QY 1205 TCCATGCTCCAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAATCCAT 1264
|||||
Db 65 TCCATGCTCCAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAATCCAT 124
|||||
QY 1265 CTTACTCATGCAACTTCCATGCAACGCAATATGTTCTCTGAAC 1311
|||||
Db 125 CTTACTCATGCAACTTCCATGCAACGCAATATGTTCTCTGAAC 171
|||||
RESULT 7
AX224396 AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
LOCATION/Qualifiers
1. 158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN

Query Match 11.1%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred. No. 5e-19;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1155 CGTGCATCTCAGATGCGATCTACATGCTGTTCAACGGTTCGTC-TTGTTCATCGTC 1213
|||||
Db 1 CGTGCATCTCAGATGCGATCTACATGCTGTTCAACGGTTCGTTGTTTCATCGTC 60
|||||
QY 1214 CAAGCTTGCTGCTATTCTGCAACCAAGAGATACCTATCCCAACAATCATCTTACTCAT 1273
|||||
Db 61 CAAGCTTGCTGCTATTCTGCAACCAAGAGATACCTATCCCAACAATCATCTTACTCAT 120
|||||
QY 1274 GCAACTTCATGCAACGCAATATGTTCTCTGAAC 1311
|||||
Db 121 GCAACTTCATGCAACGCAATATGTTCTCTGAAC 158
|||||
RESULT 8
AC147602 AC147602 186199 bp DNA linear HTG 17-AUG-2004
LOCUS Zea mays clone ZMMB0334A01, *** SEQUENCING IN PROGRESS ***, 6
DEFINITION Ordered pieces.
AC147602
AC147602.5 GI:51315585
VERSION HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 Birren, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bhatti, A.K.
and Messing, J.
Zea mays, clone ZMMB0334A01
Unpublished
2 (bases 1 to 186199)
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bhatti, A.K.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collins, A., Cook, A., Cooke, P., Corum, B., Deavellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunthang, P., Pierre, N.,
Rachupka, A., Ramasamy, V., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vasilev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zainoun, J., Zembock, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (19-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 186199)
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bhatti, A.K.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collins, A., Cook, P., Corum, B., Deavellano, K., Diaz, J.S., Dodge, S.,
Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P.,
Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L.,
Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J.,
Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Lander, E., Levine, R., Lindblad-Toh, K.,
Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P.,
Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J.,

TITLE
JOURNAL
COMMENT

Meneus L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pletie, N., Rachupke, A., Ramasamy, U., Raymond, C., Retta, R., Rhee, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V., S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zaidoun, J., Zembek, L., Zimmer, A. and Zody, M. Direct Submission
Submitted (17-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 17, 2004 this sequence version replaced gi:4658659.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@broad.mit.edu

Bharti, AK and Messing, J: The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
(<http://pgr.rutgers.edu>)
Butler, E and Wing, R: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (<http://www.genome.arizona.edu>)

Project Information
Center project name: L30003
Center clone name: 334_A1

Consensus Information
This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not in either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository
(<http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?>). An exact list of reads used in this assembly are available at
<http://www.broad.mit.edu/annotation/plants/maize/randomclones.html>.

NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 100617: contig of 100617 bp in length
* 100618 100717: gap of unknown length
* 100718 104730: contig of 4013 bp in length
* 104731 104830: gap of unknown length
* 104831 115104: contig of 10274 bp in length
* 115105 115204: gap of unknown length
* 115205 116396: contig of 41192 bp in length
* 116397 156496: gap of unknown length
* 156497 179936: contig of 23440 bp in length
* 179937 180036: gap of unknown length
* 180037 186199: contig of 6163 bp in length.

location/Qualifiers
1 .186199
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZM85BC0334A01"

ORIGIN
Query Match
Best Local Similarity 71.4%; Pred. No. 2,7e-05;
Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

840 TTTCAAAAAAATAGTATTTCTTTCTTAAATAGAAAACCTTAGAAAAATAGAG 899
162117 TTCTAGTGAATTAATTAATTAATTTCTTAAACAATAATTTTGTGTAATAAGAG 162176

Qy 900 TTGCCAGACTAGCCTAGATGTTTCCCAATTAATTACAA 940
162177 TTTCCAACTAACCCCTCAAGTATATATGAAATGACAGACA 162217

Db 162058 CTTTCTTTTACTTTTATGAGCTAGTTGGAAACCCATATT -TTCTAAGAAATTTTCAATT 162116

Qy 780 CTTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTTTCTTCAAAAGAAATTTGATT 839
162058 CTTTCTTTTACTTTTATGAGCTAGTTGGAAACCCATATT -TTCTAAGAAATTTTCAATT 162116

Db 162117 TTCTAGTGAATTAATTAATTAATTTCTTAAACAATAATTTTGTGTAATAAGAG 162176

Qy 900 TTGCCAGACTAGCCTAGATGTTTCCCAATTAATTACAA 940
162177 TTTCCAACTAACCCCTCAAGTATATATGAAATGACAGACA 162217

RESULT 9
AC117267
LOCUS
DEFINITION
Dicyostelium discoideum chromosome 2 map 5836255-5862024 strain
AX4, complete sequence.
ACCESSION
AC117267.2 GI:42733680
VERSION
AC117267.2 GI:42733680
KEYWORDS
HTG.
SOURCE
Dicyostelium discoideum
Dicyostelium discoideum
Eukaryota; Eukaryota; Dicyostelidae; Dicyostelium.
1 (bases 1 to 25769)
Gloekner, G., Eichinger, L., Szafianski, K., Pachbat, J., Dear, P., Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K., Tunggal, B., Cox, E., Quail, M.A., Platzner, M., Rosenthal, A. and Noegel, A.A.
Sequence and analysis of chromosome 2 of Dicyostelium discoideum
Nature 418 (6893), 79-85 (2002)
12097910
MEDLINE
22092622
PUBMED
REMARK
The Dicyostelium Genome Sequencing Consortium
2 (bases 1 to 25769)
Baumgart, C.
DIRECT SUBMISSION
Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 25769)
Baumgart, C.
DIRECT SUBMISSION
Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
Biology, Beutenbergstr. 11, Jena 07745, Germany
On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from GeneID may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(<http://genome.imb-jena.de/dicyostelium/>)
and the University Cologne, Institute for Biochemistry I
(<http://www.uni-koeln.de/dicyostelium/project.shtml>)
Funding
Agency: Deutsche Forschungsgemeinschaft (DFG).

location/Qualifiers
1 .25769
/organism="Dicyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
complement(join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="GeneID exon scores (in order of location ranges):
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1

[illegible]

SPSSSSSTSPSCLSSSSSSTPSQSIENKODINNNCNNTNGNPFEEBEEBKXV
BOOSIEETFEIESFEETKEVVEFPPEPEFLPTISIEBETKEYTSLSLSLSEIV
DMYDDDKSSTSSQSSSSSSSSSPSPKIKNDNNNTTTTATTTTSTKTATSS

Query Match 5.4%; Score 73.4; DB 3; Length 25769;

Best Local Similarity 49.4%; Pred. No. 0.00015;

Matches 217; Conservative 0; Mismatches 221; Indels 1; Gaps 1;

```
QY 438 CCGAACAAGATTAAAAAATTAACCAAGTAATCACTCGAAGCTATCATGTAA 497
DB 2392 CGAACAACACACTAAAAAATTTGTTGAGAAAAAATTTTAAAAA 2451
QY 498 TGTAAAGAACATCATTAATAACCAAGTCTTAAATAAAGCATATTTCGAAG 557
DB 2452 TATGAAAAAATTAATAAATTAATAAAGTCTTAAATAAATTAATCAAAA 2511
QY 558 AGACAATTAATGTACAGTTTCAAAATCTAAGAGCAAAATTAATGAAGTAA 617
DB 2512 AATTTATTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 2571
QY 618 CTATGACGTTGAGATTTTCTTTTCACTTGTATTTGTTATTTTATATACAT 677
DB 2572 ATAAAAATTAATTAATTAATTAATTAATTAATTTTATTTTATTAATGA 2630
QY 678 TTTCTCTCTTACATAGATTTTCTCGATTTTAATAATGACTATAAGTCAT 737
DB 2631 TAAATTTTAAAGATATTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2690
QY 738 TTTATTAAGAGCAGCATGTCGTAGATTTCTCGTCAAAATCTTCTGATTTT 797
DB 2691 TTTAAATTTGGGTTGATTAATAAATTTCTTTAAATTTTTCAGAAATATCT 2750
QY 798 AGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAATTA 857
DB 2751 TTTTCTTTCAAAACCAATTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2810
QY 858 TATTTCTCTTATAAAT 876
DB 2811 TTAATTTTCTTTTATTTAT 2829
```

RESULT 10

AF034389/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

SOURCE	Dicystostelium discoideum
ORGANISM	Eukaryota; Mycetozoa; Dicystostellida; Dicystostelium.
REFERENCE	1 (bases 1 to 3576)
AUTHORS	Bescallante, R., Weeseels, D., Soli, D.R. and Loomis, W.F.
TITLE	Chemotaxis to cAMP and slug migration in Dictyostelium both depend on migR, a BTB protein
JOURNAL	Mol. Biol. Cell 8 (9), 1765-1775 (1997)
PUBMED	97453232
REFERENCE	2 (bases 1 to 3576)
AUTHORS	Bescallante, R., Weeseels, D., Soli, D. and Loomis, W.F.
TITLE	Direct Submission
JOURNAL	Submitted (27-JUN-1997) Department of Biology, University of California, San Diego, 950 Gilman Drive, La Jolla, CA 92093-032, USA
FEATURES	
source	Location/Qualifiers 1..3576 /organism="Dicystostelium discoideum" /mol_type="genomic DNA" /strain="AX4" /sub_strain="AK244" /db_xref="taxon:44689" /chromosome="4" map=4.25 884..3423 /gene="migA" join(<884..2899,2998..>3423) /gene="migA" product="MigA" join(884..2899,2998..3423) /gene="migA" function="essential for slug migration" note="N-termilus of this protein is similar to other BTB domain proteins, such as the Mus musculus and human zinc finger proteins encoded by GenBank Accession Numbers U1456 and U16896, respectively"
CDS	/codon_start=1 /product="MigA" /protein_id="AA047544.1" /db_xref="GI:1841872" translation="MATNTITNNNNNNNNNNNNINTEVEFNNPFGKSLGEFKKYTDIKKGGOIPPHRIITLAASSEPFGBLLSDPKSSOSYTELKQDPFRVPMYIEPIETGIIYSERNIPLMAVDHILRDQLCLFLDONLARENVLTLKSYIDHFHTTSASSSSSSNNNYTTTITTTTTTAASSNPKNKLINCVTGQQQQQQQQQQQQQQQQQQQQQIQDFSPASIIFYOIDHDYLAVNSEYTLRYLVNFLEKPGFNFEITYTNITDOLKVSGLMSQIRYILMPYDLILRKDPLPEELIFETLLETKREFEKTSSPSSTSSPSPTNS.PRTKOTSPSKYQQQQQQQQQQQQQQQQQQQQQQQQQQQQPPQPQPPOIIOQCPQOIIOOCPQOIQOIQOAPQOPQTQMVGQOOITQOQQQQQQQQOQQOQTOQMQOOLIQQOIQOQQQQQQQQQQQOQIIKOFPSSPSSSFISLASOKLSINNATNTNNNNNNNNNNVVNTNNNIINNANNNKQTQSNNNDKI LKRQRPQPOSILFEYSFDPDFKGIIFWISTDGNEKSNHSYTSKIKITSIDDKNL LYDIIVLEPNAFMTKKDVASWMIDLGNRVTVPMYTIRGLSKDSLRTMPFOGS TNGOMTYLKHGTNDPSLNXYKXATHSWPVTCETAFRYRLLOTGRKSNRNRFVLIGG LEIYGELCEYNPNEN">
exon	<884..2899 /gene="migA" number=1 /gene="migA" 951..1289 /gene="migA" note="encodes BTB domain; similar to the Drosophila melanogaster BTB domain encoded by GenBank Accession Number U01333"
misc_feature	1253-1254 /gene="migA" note="gene interruption in mutant" replace="obsrt external plasmid, approximately 4 kb"
variation	/number=1 /gene="migA" 2900..2997 /gene="migA" number=1 /gene="migA"
intron	2998..>3423 /gene="migA"
exon	

ORIGIN	/number=2
Query Match	5.4%; Score 71.2; DB 3; Length 3576;
Best Local Similarity	47.2%; Pred. No. 0.00043;
Matches 217; Conservative	0; Mismatches 243; Indels 0; Gaps 0;
OY	412 TGTCTAGTGCACATATTGGCATCTACCCAGAACAGATTAAAAAAAATAACCAAGTAC 471
DB	852 TATTATTATTATTATTATTATTATTAACAATTTGTGTTTGTATAAAAAATAAAAAATAAAAA 793
OY	472 TAATCCACTCGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACCAAGATCCT 531
DB	792 TAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAACAACAAATTAATTTAA 733
OY	532 CTTAAAAACAGCATTTTTCGAAAGACAAATTATGTTCACGTTTACAAACATCTAAG 591
DB	732 TTTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 673
OY	592 AGCAGCAAAATTATTCGAAAGTACGATCATGACGTTTCAGATTTTCTTTTCATCTGT 651
DB	672 AAAAAAAAAAAAAAAAAAATATATATGAGTGTGACAGAAATGTTTTTTTTTTGTTG 613
OY	652 TATTTTGTATTTGTTTAAATACATTTCTCTTACAAATGAGATTTTCTTCGGA 711
DB	612 TTGTTATTTTGTATTTTAAATTAATTTTAAATTAACAATAAATACTTTAAACAAAC 553
OY	712 TTTTATTAATATGACTAATAAGTCAATTTTATTAATAAGACGCAATGTCGATTCGT 771
DB	552 ACACACACACACACAAATGTTTTTTTTTTTGTGTTTTTGTGTTTTTTTAAACCT 493
OY	772 TCAAAAACTTTTCGATTTTAAAAAGCAGTGTGGCAACCTGTTCTTTCAAGAAAT 831
DB	492 TTCAATGCGCTTTGATTTAGTTAGTTTTTTTTTTTTTTTGAATGTTTTTTTGA 433
OY	832 TTTGATTTTTCAAAAAATAATGATTTATTTTCTCTTAT 871
DB	432 ATTTTTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 393
RESULT 12	
EX957346/C	
LOCUS	EX957346 149526 bp DNA linear HTG 08-OCT-2004
DEFINITION	Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.
ACCESSION	EX957346
VERSION	EX957346.13 GI:54019944
KEYWORDS	HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Danio rerio (zebrafish)
ORGANISM	Danio rerio
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
	Cypriniformes; Cyprinidae; Danio.
	1 (bases 1 to 149526)
	Beasley,H.
	Direct Submission
	Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
	Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
	zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
	On Oct 9, 2004 this sequence version replaced gi:53850295.
REFERENCE	Genome Center
AUTHORS	Center: Wellcome Trust Sanger Institute
TITLE	Center code: SC
JOURNAL	Web site: http://www.sanger.ac.uk
	Contact: zfish-help@sanger.ac.uk
	Project Information
	Center project name: ZC117K16
	Summary Statistics
	Assembly program: XGAP4; version 4.5
	Chemistry: Dye-terminator; 100% of reads
	Consensus quality: 149526 bases at least Q40
	Consensus quality: 149526 bases at least Q30
	Consensus quality: 149526 bases at least Q20
	Insert size: 149526; sum-of-contigs
COMMENT	

Insert size: 149176; 6.5% error; agarose-fp
Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality
coverage: 12.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

Source

Location/Qualifiers
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-117K16"
/clone_11b="CHOR1-211"
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN

Query Match 5.4%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00047;
Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

QY 488 TATCATGTAAATGTTTAAAGAACATCTATTAAACACGATCTCTTAAAAACAGCAT 547
DB 109063 TTTTAAATATATGAACAAAAATCTATATTTTACTTTTGCCCTTGTTAGTCATTAGTCAT 109004
QY 548 ATTTCGAAAGACAATTTTGTTCAC---GTTTACAAACATCTAAGCGCAAAATTA 603
DB 109003 GTTTTAAAGATTTAAATAATTTTAAACATGTTTATATCAATATATATGACTAAGAGAG 108944
QY 604 TATCGAAAGGTAGCTATGACGTTGAGATTTTCTTTTCATTCGTGTTATTTGTTAT 663
DB 108943 CCAAAAGTACTAATAGCATTTTAAATGTTTATTTTCGATTTTATTTTATTTTATTT 108884
QY 664 GTTTTATATATACATTTTCTCTCTTACATATAGATGATTTTCTTCGATTTTAAATG 723
DB 108883 AT 108824
QY 724 ACTATATAGCATTTTATATATAGAGACGACGATGCTGATGATTCGTTCAAAATCTTT 783
DB 108823 TATATATTTTAAATTTATATATATATATATATATATATATATATATATATATAT 108764
QY 784 CTGATTTTAAAGTAGTGTGGCAACCGCTGTTCTTCAAGAAATTTGATTTTTC 843
DB 108763 ATTATATCATATTTAAACCTTATATATGTTATATATATATATATATATATATATAT 108704
QY 844 AAAAAAATAGTTTATTTTCTTTATATATATATATATATATATATATATATATATATATAT 903
DB 108703 AGAAAAACAAAGTAT 108644
QY 904 CAGACTAGCCCTAGAAATGTTTCCCAATATATATATATATATATATATATATATATATAT 963
DB 108643 TATATATGCCCCCTTATAGCATCTTATATATATATATATATATATATATATATAT 108584
QY 964 AGCCCATATATATTTTAAACGAAACTGAATCGAGCAACCAATCGACCTATTC 1023
DB 108583 ACAT 108524
QY 1024 TCTAGATTAGTAAAA 1039
DB 108523 GTTTTATATAGTAAAA 108508

RESULT 13
PFMAL4P1/c PFMAL4P1 347582 bp DNA linear INV 29-JAN-2003
DEFINITION Plasmodium falciparum MAL4P1.
ACCESSION AL034557 AL844503

VERSION

AL034557.8 GI:23498126

KEYWORDS

Plasmodium falciparum 3D7

SOURCE

Plasmodium falciparum 3D7

ORGANISM

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE

1

Hall, N., Patn, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barton, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagers, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Laike, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moul, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tiley, A., Unwin, L., Whitehead, S., Woodward, D., Suleston, J.E., Craig, A., Newbold, C. and Barrell, B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
Nature 419 (6906), 527-531 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

JOURNAL

COMMENT

FEATURES

source

1. 347582
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P1"
10384. 32958
/note="rep20 Repeats"
join(35153. 41725, 42757. 44124)
/gene="VAR"
/note="synonym: PF00005"
join(35153. 41725, 42757. 44124)
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific surface protein var-2 SMALL-026033 (EMBL:L40609) (2664 aa) fasta scores: E(): 7.8e-149, 43.1% id in 2872 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PEEMP1)"
/protein_id="CAD49094.1"
/db_xref="GI:23498127"
/db_xref="GOA:O81220"
/db_xref="UniProt/TREMBL:O81220"
/translation="MVRGSGSGGVDVDGDIHQAKHLIDSIGKKYHGVKNGADCT GNSGDKNTIDLDKEDLOKAPNINFLIGTDPCQVEDYINNHVNGDQKGRFYCTE LSGKRFNPFSDTLGGQCTNSKMRSGBCACPYRRLLHCHNLDSLETTSTYADTL LLEVCAAAKYEGQSINTHTYTKHSHNKSPOGLCVLARSDIDIVAGKDLFYGAT YESAREKLENGLKEVFGKIHGLSEBAKKTQODGNYVQLRBDMTYANRETVKAI TCEVSGNNYPRATCGDEKNPGLTSKQKCRCDKXGKPKSGGNNYIVPTYPDYPOY LWEREMADPCRLRKHKLKDAIKCRKNGEKYCDLNRDKXTYASGKYAPFPEDE CKDCOYSCAPFDVMDNQKLEFLKQKRYTKEITSGSGSGSGSKRKRDRATTTNYSYE KKFYELKGTXTKVNNPFLKLNDEVCITKNDIDGNAIDPKVNHSSAKKGDANK TPTRTKYCEACPCWGAKEVGGWKAERKSCQTDYDIDPKTTTITLIGDTRKSDMWK KYKFCNGNGNGEKSATPNATSRKGGKQDMQEMKICVYDENKEKTKGSDAINTCVL ODGKHREOKVTSYNAPFMKVDHMLDSVWERELNSCINNAKSONCKNKKCKE CCDFEKWKQKKKEKEMEAIKOHFGKQXIIIDTGCDAVITLAAYLKLEIANDTBEKS


```

Db      335424  TTTTCAGTATTAAGATTTTAACGAAATATGAGAAATGAAAGAAATTTGGATG 335365
Qy      744    TAAAGACAGCATGTCGTAGATTTCTGTTCAAAAATCTTTCTGATTTTTTTAAGAGCTAG 803
Db      335364  TAAATTTTAAATTAATAAATAAAGAAATATATATTTTATTAATTAATTAATAAAG 335305
Qy      804    TTGGCAACCTGTTCTTCCAAAGAAATTTGATTTTTTCAAAAAAATTTAGTTATTT 863
Db      335304  TTTATTTTATTTATATATATATTTTATTTTAAATTTTAAATTCATATATTTTAAATTT 335245
Qy      864    CTCCTTATTAATAAGAAACCTTAGAAAAATAGAGTTGCCAGACTGACCTGAAATGTT 923
Db      335244  ATATTTTTTTTTTTTTTTTGGCTTATATATATATATATATACGTATATATATATATTT 335185
Qy      924    TTCCCAATAAATTAACAATCACTGTATTAATTTTGGCAGCCCAATAATTTATTTAA 983
Db      335184  TTTTTTTTTTTTTTTTAAATTAATATTTATTTTATTTTCTACATATATATAATATA 335125
Qy      984    CCGAACTGAAATCGAGCGAAACCAATCTGAGCTATTTCTTATGATTAGTAAAGGCA 1043
Db      335124  TAAATTTTATTTTAAAGTGATATGTTATTTTAAATTTTAAATTTAAATTTAGT 335065
Qy      1044  GAGAGAGAGAGAGAA 1059
Db      335064  TATCTATATATTAATAA 335049

```

RESULT 14

```

LOCUS      CO422641      883 bp      DNA      linear      PAT 28-JAN-2004
DEFINITION Sequence 7675 from Patent WO0151628.
ACCESSION  CO422641
VERSION     CO422641.1  GI:41374870
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1. 883
AUTHORS     Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE       Novel gene, compositions, kits, and methods for identification,
JOURNAL     assessment, prevention, and therapy of breast cancer
            Patent: WO 0151628-A 7675 19-JUN-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES
            Location/Qualifiers
            source          1..883
                           /organism="Homo sapiens"
                           /mol_type="unassigned DNA"
                           /db_xref="taxon:9606"

```

ORIGIN

```

Query Match      5.4%; Score 70.2; DB 6; Length 883;
Best Local Similarity 39.2%; Pred. No. 0.0007;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;

Qy      419  TGTCCACTATGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAAGTAATCCA 478
Db      873  TGTNNNNNNNANACCTTAAANATTTNNNANNTNANNAANAAAAATATTTAAAN 814
Qy      479  CTCGAAGCATATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCCTTAAA 538
Db      813  TANAATTTTAAAAAATTAATTTNANNTNANNTNANNTTNTTAAAAAANNNAAAAA 754
Qy      539  AACAGCATATTTGCAAAAGAGACAAATATATGTTACAAACATCTAGAGGACGA 598
Db      753  TTAANNNTTTNTNANTATTAACCAAAATTTTTTTTAAABAAAATTTTTTAAANNT 654
Qy      599  AATTATATCGAAAGTAAGCTATGAGCTTGAATTTCTTTTTCATTTCTGTTATTTG 658
Db      693  AATTAATTTNAAATTTTTTTTATTAATAAATAAATTTTAAATTTTANCAANTNTTTT 634
Qy      659  TTATTTGTTTATATACATTTTCTCTCTTACATAGAGATTTTCTTCGATTTTATA 718

```

```

Db      633  TTTNTTNNATTAATAAATTTTAAATTAATAAANANNTTTTTTTATTAANATTTAA 574
Qy      719  AATGACTATTAAGCATTTTATATATAGACAGCATGTCGTAGATTCGTTCAAAAA 778
Db      573  AAAAAANNTTTTTTTTTTATANTTTATTAATAAATAATTTATTTNTTTTCTNAAAA 514
Qy      779  TCTTTGATTTTATTAAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
Db      513  AAAAAAATTTTNNAAATNTTTTAAAAACCTTNAATTAANANAAATTTTNTNTT 454
Qy      839  TTTTCAAAAAAATAGATTATTTTCTTTTAAATTAATAAGAAACCTTAGAAAAATAG 898
Db      453  NNNATTTAAAAATTTATTTTATTTTATNTTCTATTAATTAANANANNAATTAAT 394
Qy      899  GTTGCAGACTAGCCCTAGAAATGTTTCCCAATAATTAACAATCACTGTATTAATTT 958
Db      393  TTTNAAAAAATATATTTTATANTTATNNANANNAATTTAAANTNANANAAAAATTT 334
Qy      959  TGGCCAGCCCAATAATTTTAAACCGAAACTGAAATCGAGGAAACCAATCTGAGCT 1018
Db      333  NTTAATTAATAAATAATTAATTTAAACNTCATTTTAAATTTATATTAATTAATAA 274
Qy      1019  ATTT 1022
Db      273  NTT 270

```

RESULT 15

```

LOCUS      BV119878/c      241 bp      DNA      linear      STS 19-MAR-2004
DEFINITION PZA01377 Zea mays CML247 Zea mays STS genomic, sequence
ACCESSION  BV119878
VERSION     BV119878.1  GI:45589251
KEYWORDS
SOURCE      Zea mays
ORGANISM    Zea mays
REFERENCE   1. 241
AUTHORS     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 241)
            McMillen, M.D., Vrch Bi, I., Schroeder, S.S. and Gaut, B.S.
            MP2-UCI Joint SNP Discovery
            Unpublished (2003)
COMMENT

```

```

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTCTACGTCGAATGCAATGCAAGA
Primer B: AGGAACCAAGCTTATGAGGGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
sequencing ready reaction with amplified DNA polymerase FS
Sequence ran on ABI 3700 sequencer.

```

```

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine kit (ABI)

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 634.708 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311

Sequence: 1 ccacggcgcctcctacgaaaa.....cgcacatcgtcttcctgaac 1311

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.16Dec04:*
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20018:*
6: geneseqn20028:*
7: geneseqn20028:*
8: geneseqn20038:*
9: geneseqn20038:*
10: geneseqn20038:*
11: geneseqn20038:*
12: geneseqn20048:*
13: geneseqn20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1311	100.0	1394	2	AAx07408	Aax07408 Zea mays
2	1311	100.0	1394	2	AAx07409	Aax07409 Zea mays
3	1311	100.0	1394	5	AAH76332	Aah76332 Z. mays M
4	1311	100.0	1394	5	AAH76333	Aah76333 Z. mays M
5	160.6	12.3	255	5	AAH76340	Aah76340 Z. mays M
6	146	11.1	158	5	AAH76334	Aah76334 Z. mays M
7	70.2	5.4	883	4	AA15210	Aa15210 Human bre
8	70.2	5.4	960	11	ACN85231	Acn85231 Breast ca
9	69.8	5.3	13400	13	ADR04296	Adr04296 Corn FT h
10	67.2	5.1	2657	2	AAZ10551	Aaz10551 DNA seque
11	66.4	5.1	6027	2	AAx58751	Aax58751 Maize dul
12	66.4	5.1	6027	8	ABX09935	Abx09935 DNA encod
13	66.4	5.1	6027	12	ADK12106	Adk12106 cDNA encd
14	66.2	5.0	439	8	ABX35844	Abx35844 Bovine ES
15	65.6	5.0	346	4	AA187279	Aa187279 Human pol
16	65.6	5.0	8056	8	ABZ10246	Abz10246 Haematopo
17	64.6	4.9	9859	13	ADs89714	Ads89714 Oligonuc1
18	64.4	4.9	9859	13	ADs89440	Ads89440 Oligonuc1
19	64	4.9	8056	8	ABZ10100	Abz10100 Haematopo
20	63.8	4.9	431	8	ABX44556	Abx44556 Bovine ES

C	21	63.6	4.9	335913	5	AA161371	Aa161371 Soybean 2
C	22	63.6	4.9	335913	5	AA161372	Aa161372 Soybean 2
C	23	63.4	4.8	12368	4	AA545494	Aa545494 Chemicall
C	24	63.4	4.8	12368	4	AA546779	Aa546779 Tumour su
C	25	63.4	4.8	12368	6	AB134085	Ab134085 Human imm
C	26	63.4	4.8	12368	6	ABK28424	Abk28424 DNA trans
C	27	63.4	4.8	12368	6	ABN80285	Abn80285 Human che
C	28	63	4.8	7624	6	AB134113	Ab134113 Human imm
C	29	62.4	4.8	10762	6	AB192322	Ab192322 Chemicall
C	30	61.4	4.7	9652	6	AB132909	Ab132909 Human imm
C	31	60.6	4.6	565	13	ACN56344	Acn56344 Cotton an
C	32	60.6	4.6	5822	6	AB133097	Ab133097 Human imm
C	33	60.4	4.6	500	5	AD173089	Ad173089 Human ova
C	34	60.4	4.6	500	5	AD173089	Ad173089 Human ova
C	35	60.2	4.6	5818	4	AA546636	Aa546636 Tumour su
C	36	60.2	4.6	8896	6	ABN80326	Abn80326 Human che
C	37	59.4	4.5	461	5	AD143834	Ad143834 Human ova
C	38	59.4	4.5	14919	4	AA546506	Aa546506 Tumour su
C	39	59.2	4.5	19459	6	ABK31213	Abk31213 Signal tr
C	40	59.2	4.5	19459	6	AB170528	Ab170528 Chemicall
C	41	59	4.5	517	13	ACN56273	Acn56273 Cotton an
C	42	59	4.5	5303	6	AB132871	Ab132871 Human imm
C	43	58.8	4.5	393	8	ABX39417	Abx39417 Bovine ES
C	44	58.6	4.5	392	5	AD143653	Ad143653 Human ova
C	45	58.6	4.5	529	13	ACN56029	Acn56029 Cotton an

ALIGNMENTS

RESULT 1
AAx07408
ID AAx07408 standard; DNA; 1394 BP.
XX
XX
AC AAx07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays M45 male tissue-preferred regulatory region.
XX
XX M45; male; tissue-preferred; regulatory region; plant cells;
KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX
PN W09859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX WPI; 1999-105628/09.
PT New nucleic acid encoding a M45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 2; Page 22-23; 39pp; English.
XX
XX The sequence is that encoding an M45 male tissue-preferred regulatory
XX region. It may be used in the construction of a vector for a method of
XX producing exogenous genes in a male tissue-preferred manner, which is
XX useful in restoring or conferring fertility, such as in hybrid seed
XX production. In conferring fertility, a monocot/dicot plant is transformed
XX with the exogenous nucleotide sequence (a male sterility gene, preferably
XX M45), which encodes a product selected from auxins, rolgs and diptheria
XX toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
XX and infertile plants

XX	Sequence
SQ	1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match	100.0%	Score 1311	DB 2	Length 1394
Best Local Similarity	100.0%	Pred. No. 1e-269		
Matches 1311	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY	1	CCATGAGTCTCTATGAAAAGAAGATGATCAATGTCCTATATCCGTTTCTTGAAGGCC	60
Db	1	CCATGAGTCTCTATGAAAAGAAGATGATCAATGTCCTATATCCGTTTCTTGAAGGCC	60
OY	61	CTTCTTCGCTTATTAAGCTGATCGGGTTTCAAAAACTTCCAGGGTGCATGAT	120
Db	61	CTTCTTCGCTTATTAAGCTGATCGGGTTTCAAAAACTTCCAGGGTGCATGAT	120
OY	121	CTCCATGTTCCACTTCTCCCACTCGCGTTGACATTTCTTGAGTGTGGTGTCCAT	180
Db	121	CTCCATGTTCCACTTCTCCCACTCGCGTTGACATTTCTTGAGTGTGGTGTCCAT	180
OY	181	CTGACCGAGGCCATCAGACACCTTTCGGGACACCCATCAAGGGGCTTTGGATGGCCCA	240
Db	181	CTGACCGAGGCCATCAGACACCTTTCGGGACACCCATCAAGGGGCTTTGGATGGCCCA	240
OY	241	CGAGACGTATCGGGTGTGTGTGATCCAGGGGATATATGTCCCCCAATCGTCACTATA	300
Db	241	CGAGACGTATCGGGTGTGTGTGATCCAGGGGATATATGTCCCCCAATCGTCACTATA	300
OY	301	TTATATATCTTTAGATATTTATTTATTTTGGAAAAATPAACAACCTTATCTTTTGTTA	360
Db	301	TTATATATCTTTAGATATTTATTTATTTTGGAAAAATPAACAACCTTATCTTTTGTTA	360
OY	361	GGGCGCTCAGATAGATTTTTCGTTTAGGGCCAGAAATCGAGACACAGCATGTCTAGTG	420
Db	361	GGGCGCTCAGATAGATTTTTCGTTTAGGGCCAGAAATCGAGACACAGCATGTCTAGTG	420
OY	421	TCCACTATTTGGACATCCAGAAACAAGTTTAAAAAATTAACCAAGTAACTAATCCACT	480
Db	421	TCCACTATTTGGACATCCAGAAACAAGTTTAAAAAATTAACCAAGTAACTAATCCACT	480
OY	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTTTAAAAAA	540
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAAAACAAGATCCTTTAAAAAA	540
OY	541	CAAGATATTTTCGAAAGACAAATTATGTTTACAGTTTACAAACATCTPAAGAGGACAAA	600
Db	541	CAAGATATTTTCGAAAGACAAATTATGTTTACAGTTTACAAACATCTPAAGAGGACAAA	600
OY	601	TTATATCGAAAGTAAAGCTATGACGTTTCAGATTTTTCCTTTTCAATCTGTGTAATTTGTT	660
Db	601	TTATATCGAAAGTAAAGCTATGACGTTTCAGATTTTTCCTTTTCAATCTGTGTAATTTGTT	660
OY	661	ATTGTTTTTATATACATTTTCTCTCTTACATAGAGTATTTTCTTCGATTTTATATAA	720
Db	661	ATTGTTTTTATATACATTTTCTCTCTTACATAGAGTATTTTCTTCGATTTTATATAA	720
OY	721	ATGACTATAAAGTCAATTTTATATATAGAGACGCAATGTCGTATGATCTTCGTTCAAAAATC	780
Db	721	ATGACTATAAAGTCAATTTTATATATAGAGACGCAATGTCGTATGATCTTCGTTCAAAAATC	780
OY	781	TTTCTGATTTTTTAAAGACTAGTTTGGCAACCCGTGTTCTTTTCAAAAGATTTGATTTT	840
Db	781	TTTCTGATTTTTTAAAGACTAGTTTGGCAACCCGTGTTCTTTTCAAAAGATTTGATTTT	840
OY	841	TTTCAAAAAAATATGATTTATTTTCTCTTTATATAAATAGAAAAACATTAGAAAAATAGAGT	900
Db	841	TTTCAAAAAAATATGATTTATTTTCTCTTTATATAAATAGAAAAACATTAGAAAAATAGAGT	900
OY	901	TGCGAGACTAGCCCTAGAGATGTTTTCCCAATATAATTAACATCAGTGTATATAATTTTG	960
Db	901	TGCGAGACTAGCCCTAGAGATGTTTTCCCAATATAATTAACATCAGTGTATATAATTTTG	960
OY	961	GCCAGCCCATTAATATTTTAAACCGAAACTGAAATCGAGGAAACCAATCTGAGCTAT	1020

Db	961	GCCAGCCCCCAATTAATTATTAAACCGAAACTGAAATGAGCGAAACCAAACTCTGAGCTAT	1020
OY	1021	TTCTCTAGATTATGTAATAAAGGAGAGAGAGAGAAGAAATCAGTTTTAAGCATTTGTC	1080
Db	1021	TTCTCTAGATTATGTAATAAAGGAGAGAGAGAGAGAAGAAATCAGTTTTAAGCATTTGTC	1080
OY	1081	TGAGATGAGCGCGTTGGCAACGATAGACACCGTATATATAGCTATAGGTGCGTACGTCA	1140
Db	1081	TGAGATGAGCGCGTTGGCAACGATAGACACCGTATATATAGCTATAGGTGCGTACGTCA	1140
OY	1141	GGTTCCGACGCTCGTGTCACTCCACATGACATACATGCTTTGTTCAACCGTTTCGTC	1200
Db	1141	GGTTCCGACGCTCGTGTCACTCCACATGACATACATGCTTTGTTCAACCGTTTCGTC	1200
OY	1201	TTGTTCCATCGTCCCAAGCGTTGCCCTATTTCTGAAACCAAGAGATATCCATCCCAAACAAT	1260
Db	1201	TTGTTCCATCGTCCCAAGCGTTGCCCTATTTCTGAAACCAAGAGATATCCATCCCAAACAAT	1260
OY	1261	CCATCTTAATCTATGCAACTTCATGACAAACAGCACATATGTTCTCTGAAC	1311
Db	1261	CCATCTTAATCTATGCAACTTCATGACAAACAGCACATATGTTCTCTGAAC	1311

RESULT 2
AA07409
ID AA07409 standard; DNA, 1394 BP.
AC
XX AAX07409;
XX
DT 08-JUN-1999 (first entry)
DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
XX Ms45; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
XX MO9859061-A1.
XX
XX PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
PI Albersen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region.
PT useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 3; Page 23-24; 39pp; English.

The sequence is that encoding an Ms45 male tissue-preferred regulatory CC region. It may be used in the construction of a vector for a method of CC producing exogenous genes in a male tissue-preferred manner, which is CC useful in restoring or conferring fertility, such as in hybrid seed CC production. In conferring fertility, a monocol/dicot plant is transformed CC with the exogenous nucleotide sequence (a male sterility gene, preferably CC Ma5), which encodes a product selected from auxins, rolB and diptheria CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile CC and infertile plants

Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match	100.0%;	Score 1311;	DB 2;	Length 1394;
Best Local Similarity	100.0%;	Pred. No. 1e-269;		
Matches 1311; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	CCATGGTGTCTCTAAGAAAAGATGAGTCAATNGTCTAATCCGTTTTCTTAGGGTCC	60
Db	1	CGATGGTGTCTTAAGAAAAGTAGTACAAATGTCTAATCCGTTTTCTTAGGGTCC	60
OY	61	CTTCTTCTGCTTATTTACTGATCGGAGTACCAAAAACTTCCACGGGTGATGAT	120
Db	61	CTTCTTCTGCTTATTTACTGATCGGAGTACGAGGAGTACCAAAAACTTCCACGGGTGATGAT	120
OY	121	CTTCATGTTCCACTTCTCCACCTGCGGTTGCAATTTCTTGATGTGCGTGGTTCCAT	180
Db	121	CTTCATGTTCCACTTCTCCACCTGCGGTTGCAATTTCTTGATGTGCGTGGTTCCAT	180
OY	181	CTGACCGAGGGCCCATCGACACCTTTCCGGGACACCATCAAGGGCTTTCCGATGCGCCA	240
Db	181	CTGACCGAGGGCCCATCGACACCTTTCCGGGACACCATCAAGGGCTTTCCGATGCGCCA	240
OY	241	CGAGACGTATCGGGGTGTGTGATCCAGGGAGTATATGTCCGCCCAATTCGCACCTATA	300
Db	241	CGAGACGTATCGGGGTGTGTGATCCAGGGAGTATATGTCCGCCCAATTCGCACCTATA	300
OY	301	TTATTATTCTTTAGATATTATTAAATTTTGGAAAAATTAACAACTTATCTTTTGTGTA	360
Db	301	TTATTATTCTTTAGATATTATTAAATTTTGGAAAAATTAACAACTTATCTTTTGTGTA	360
OY	361	GGGCGCTCAGACATAGATTTTCGCTTAGGGCCCGAAATGCGAGGACACAGCATGTCTAGTG	420
Db	361	GGGCGCTCAGACATAGATTTTCGCTTAGGGCCCGAAATGCGAGGACACAGCATGTCTAGTG	420
OY	421	TCACATATTGGCAGTACCCAGAAACGAATTTAAAAAATAACCAAGTACATAACCACT	480
Db	421	TCACATATTGGCAGTACCCAGAAACGAATTTAAAAAATAACCAAGTACATAACCACT	480
OY	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATAAACCAAGATCCTTTAAAAAA	540
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATAAACCAAGATCCTTTAAAAAA	540
OY	541	CAGCATATTTCGAAAGAGACAATAATGTTACAGTTTACAAATCTTAAGAGGACAA	600
Db	541	CAGCATATTTCGAAAGAGACAATAATGTTACAGTTTACAAATCTTAAGAGGACAA	600
OY	601	TTATATCGAAAGGTAGAGCTATGAGTTCAGATTTTCTTTTCATCTGTGTAATTTGTT	660
Db	601	TTATATCGAAAGGTAGAGCTATGAGTTCAGATTTTCTTTTCATCTGTGTAATTTGTT	660
OY	661	ATTGTTTTTATATACATTTTCTTCTTACATAGAGATTTTCTCCGATTTTATATA	720
Db	661	ATTGTTTTTATATACATTTTCTTCTTACATAGAGATTTTCTCCGATTTTATATA	720
OY	721	ATGACTATTAAGTCAATTTTATATAAGACGCAATGTCTGATGATCTGTCAAAATC	780
Db	721	ATGACTATTAAGTCAATTTTATATAAGACGCAATGTCTGATGATCTGTCAAAATC	780
OY	781	TTTTCGATTTTTTTAAGAGCTAGTTTGGCAACCCGTTCCTTCAAAATTTTGATTTT	840
Db	781	TTTTCGATTTTTTTAAGAGCTAGTTTGGCAACCCGTTCCTTCAAAATTTTGATTTT	840
OY	841	TTCAAAAAAATTAGTTTATTTTCTCTTATTAATAATAGAAAACACTTAGAAAAATAGT	900
Db	841	TTCAAAAAAATTAGTTTATTTTCTCTTATTAATAATAGAAAACACTTAGAAAAATAGT	900
OY	901	TGCGAGCTAGCCCTAGAAATGTTTTCCCAATTAATTACATCACTGTGTAATTAATTG	960
Db	901	TGCGAGCTAGCCCTAGAAATGTTTTCCCAATTAATTACATCACTGTGTAATTAATTG	960
OY	961	GCCAGCCCTATTAATTTATTAACCGAAATCGAAATCGAGGAAACCAATCTGAGCTAT	1020
Db	961	GCCAGCCCTATTAATTTATTTAAACCGAAATCGAAATCGAGGAAACCAATCTGAGCTAT	1020
OY	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCATTTGCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCAGTTTAAAGTCATTTGCC	1080
OY	1081	TGAGATGTGCGGTTGGCAACGATGCCACCGTATCTATAGCTCATAGTCCCTAACGTCA	1140

[illegible]


```

Db      61 CTTCCTTCGCTTATTACTGACTGAATCGGGGTTACAAAACTTCACGGGTGCATGAT 120
Qy      121 CTCACATGTTCCACTTCTCCCACTCGGTGGCACTTCTTGATGTGGTGGTTCCCAT 180
Db      121 CTCACATGTTCCACTTCTCCCACTCGGTGGCACTTCTTGATGTGGTGGTTCCCAT 180
Qy      181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCCATCAAGGGCTTTCGATGGCCCA 240
Db      181 CTGACCGAGGCCCATGAGACACCTTTCGGGACACCCATCAAGGGCTTTCGATGGCCCA 240
Qy      241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACTTAT 300
Db      241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACTTAT 300
Qy      301 TTATTTATCTTTAGATATTTATTTTGGAAAAATACAACTTATATCTTTGTGTA 360
Db      301 TTATTTATCTTTAGATATTTATTTTGGAAAAATACAACTTATATCTTTGTGTA 360
Qy      361 GGGCTCAGCATGATTTTCGTTAAGGCCCAGAAAAATGCGAGGACGACCATGTCTAGT 420
Db      361 GGGCTCAGCATGATTTTCGTTAAGGCCCAGAAAAATGCGAGGACGACCATGTCTAGT 420
Qy      421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATACCAAGTAACTATCCACT 480
Db      421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATACCAAGTAACTATCCACT 480
Qy      481 CGAAGACTATCATGTATGTTTAAAGAAACATCTATTTAAACGACATCCTCTTAAAAA 540
Db      481 CGAAGACTATCATGTATGTTTAAAGAAACATCTATTTAAACGACATCCTCTTAAAAA 540
Qy      541 CAAGCATATTTGCAAAAGACAAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
Db      541 CAAGCATATTTGCAAAAGACAAATATGTTACAGTTTACAAACATCTAAGAGCGACAA 600
Qy      601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCTTTCTGTTATTTGTT 660
Db      601 TTATATCGAAAGGTAAAGCTATGACGTTTCAGATTTTCTTTTCTTTCTGTTATTTGTT 660
Qy      661 ATGTGTTTATATACATTTCTTCTTCAATAGATGATTTCTTCGATTTTATTA 720
Db      661 ATGTGTTTATATACATTTCTTCTTCAATAGATGATTTCTTCGATTTTATTA 720
Qy      721 ATGACTATTAAGTCAATTTTATATTAAGAGCAGCATGTCTGTTCTGTTCAAAAAATC 780
Db      721 ATGACTATTAAGTCAATTTTATATTAAGAGCAGCATGTCTGTTCTGTTCAAAAAATC 780
Qy      781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Db      781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Qy      841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAAACACTTAGAAAAATAGAGT 900
Db      841 TTCAAAAAAATAGTTTATTTTCTCTTATTAATAAGAAAAACACTTAGAAAAATAGAGT 900
Qy      901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTACAATCACTGTATATAATTTTG 960
Db      901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTACAATCACTGTATATAATTTTG 960
Qy      961 GCGAGCCCCATTAATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Db      961 GCGAGCCCCATTAATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Qy      1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTAAAGTCACTTGTCC 1080
Db      1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAAATCAGTTTAAAGTCACTTGTCC 1080
Qy      1081 TGAAGATGTGCGGTTTGGCAACGATAGCCACCGTAAATCATAGCTCATAGGTGCTTACGTCA 1140
Db      1081 TGAAGATGTGCGGTTTGGCAACGATAGCCACCGTAAATCATAGGTGCTTACGTCA 1140
Qy      1141 GGTTCGGCAGCTTCTGTTTCATCTCAATGAGCATTAAGTGTGTTCAACCGTTTCGTC 1200
Db      1141 GGTTCGGCAGCTTCTGTTTCATCTCAATGAGCATTAAGTGTGTTCAACCGTTTCGTC 1200

```

```

Qy      1201 TTGTTCCATGCTGCAAGCCTTGCTTATCTGTAACCAAGAGATTACTACTCCCAACAAT 1260
Db      1201 TTGTTCCATGCTGCAAGCCTTGCTTATCTGTAACCAAGAGATTACTACTCCCAACAAT 1260
Qy      1261 CCATCTTACTATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311
Db      1261 CCATCTTACTATGCAACTTTCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 4
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
  hybrid seed; ds.
XX
OS Zea mays.
XX
PN MO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
  essential for initiating transcription of the Ms45 gene useful for
  mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
  comprising nucleotide sequences essential for initiating transcription of
  the Ms45 gene. A method of mediating male fertility in a plant is
  provided that involves introducing an expression vector comprising a
  promoter operably linked to (I) into a plant where the exogenous gene
  impacts male fertility of the plant and (II) controls expression of the
  exogenous gene. A method of producing hybrid seeds is also provided. The
  present invention represents a nucleic acid sequence encoding an Ms45 male
  tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match      100.0%; Score 1311; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1e-269;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CCATGCTGCTCTATGAAAAAGATAGTACATATGTCATATCCGTTTCTTAGGCTCC 60
Db      1 CCATGCTGCTCTATGAAAAAGATAGTACATATGTCATATCCGTTTCTTAGGCTCC 60
Qy      61 CTTCCTTCGCTTATTACTGACTGAATCGGGGTTACAAAAAATTCACAGGGTGCATGAT 120
Db      61 CTTCCTTCGCTTATTACTGACTGAATCGGGGTTACAAAAAATTCACAGGGTGCATGAT 120
Qy      121 CTCACATGTTCCACTTCTCCCACTCGGTGGCACTTCTTGATGTGGTGGTTCCCAT 180
Db      121 CTCACATGTTCCACTTCTCCCACTCGGTGGCACTTCTTGATGTGGTGGTTCCCAT 180

```


QY 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCATCAAGGGCTTTGGGATGGCCA 240
 DB 181 CTGACCGAGGCCCATCAGACACCTTTGGGACACCATCAAGGGCTTTGGGATGGCCA 240
 QY 241 CGAGACGTATCGGGTGTGTGATCGAGGGGATATATGTCCCAATCGTCACTATA 300
 DB 241 CGAGACGTATCGGGTGTGTGATCGAGGGGATATATGTCCCAATCGTCACTATA 300
 QY 301 TTATATATCTTTAGATATATATTTTAAAGAAAAATTAACAATCTTATCTTTGTGA 360
 DB 301 TTATATATCTTTAGATATATATTTTAAAGAAAAATTAACAATCTTATCTTTGTGA 360
 QY 361 GGGCCCTCAGCATAGATTTTGGCTTAAAGGGCCAGAAAATCGAGGACGACCATGCTAGTG 420
 DB 361 GGGCCCTCAGCATAGATTTTGGCTTAAAGGGCCAGAAAATCGAGGACGACCATGCTAGTG 420
 QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480
 DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480
 QY 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTAACTATCCACT 540
 DB 481 CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACCAAGTAACTATCCACT 540
 QY 541 CAAGCATATTTCCAGAAAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA 600
 DB 541 CAAGCATATTTCCAGAAAGACAAATTTATGTTACAGTTTACAAACATCTAAGAGGACAAA 600
 QY 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTTATTTTGTG 660
 DB 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTTATTTTGTG 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTCTTAAACATAGAGATTTTCTTCCGATTTTATATA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTCTTAAACATAGAGATTTTCTTCCGATTTTATATA 720
 QY 721 ATGACTATAAAGTCATTTTATATTAAGAGACGAGTGTGTAGATCTCGTTCAAAAATC 780
 DB 721 ATGACTATAAAGTCATTTTATATTAAGAGACGAGTGTGTAGATCTCGTTCAAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTTGAATTT 840
 DB 781 TTTCTGATTTTTTAAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAAATTTTGAATTT 840
 QY 841 TTTCAAAAAAATTAAGTTATTTTCTCTTAAATTAAGAAACATCTTGAAGAAATAGAGT 900
 DB 841 TTTCAAAAAAATTAAGTTATTTTCTCTTAAATTAAGAAACATCTTGAAGAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGTATTAATTTTG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGTATTAATTTTG 960
 QY 961 GCCAGGCCCATTAATTTTAAACCGAACTGAAGTGAAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGGCCCATTAATTTTAAACCGAACTGAAGTGAAGCGAAACCAATCTGAGCTAT 1020
 QY 1021 TTTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAACTCATTTGCC 1080
 DB 1021 TTTCTTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCAAGTTTAACTCATTTGCC 1080
 QY 1081 TGAGATGTGGGATTTGGCAAGATAGCAGCGTAATATAGCTCATAGGGGCTTAGGTC 1140
 DB 1081 TGAGATGTGGGATTTGGCAAGATAGCAGCGTAATATAGCTCATAGGGGCTTAGGTC 1140
 QY 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCATACTACATGCTTTGTTCAACGTTGCTC 1200
 DB 1141 GGTTCGGAGCTCTCGTGTATCTCAATGAGCATACTACATGCTTTGTTCAACGTTGCTC 1200
 QY 1201 TTGTTTCATGCTGCAAGCCTTGGCTATTTCTGAACCAAGAGATTAAGTCTCCAAACAT 1260
 DB 1201 TTGTTTCATGCTGCAAGCCTTGGCTATTTCTGAACCAAGAGATTAAGTCTCCAAACAT 1260
 QY 1261 CCATCTTACTCATGCAACTTCATGCAACGCAACGATATGTTTCCGAAC 1311

DB 1261 CCATCTTACTCATGCAACTTCATGCAACGCAACGATATGTTTCCGAAC 1311
 RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.
 XX
 AC AAH76340;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 promoter fragment.
 XX
 KM Ms45; male tissue; regulatory region; transcription; male fertility;
 XX
 OS hybrid seed; promoter; ds.
 XX
 PN Zea mays.
 XX
 PN W0200160997-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US04527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Example 5; Fig 8; 50bp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (1)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (1) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (1) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SO Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
 Query Match 12.3%; Score 160.6; DB 5; Length 255;
 Best Local Similarity 97.6%; Pred. No. 1.3e-24;
 Matches 163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1145 CGGCACTCTCGTGTATCTCAATGAGCATACTACATGCTTTGTTCAACCGTTGCTTGT 1204
 DB 5 CCGGATGCCCGTGTATCTCAATGAGCATACTACATGCTTTGTTCAACCGTTGCTTGT 64
 QY 1205 TCCATGTCCAAAGCCTTGGCTATTTGAAACCAAGAGATTAAGTCTCCAAACATTCAT 1264
 DB 65 TCCATGTCCAAAGCCTTGGCTATTTGAAACCAAGAGATTAAGTCTCCAAACATTCAT 124
 QY 1265 CTTACTCATGCAACTTCATGCAACGCAACGATATGTTTCCGAAC 1311
 DB 125 CTTACTCATGCAACTTCATGCAACGCAACGATATGTTTCCGAAC 171
 RESULT 6
 AAH76334
 ID AAH76334 standard; DNA; 158 BP.
 XX
 AC AAH76334;
 XX

DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; db.
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX
Query Match 11.1%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 1.5e-21;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
QY 1155 CGTGCATCTCACATGCACTACTACATGCTTGTTCACCGTTCGTC-TTGTTCATCGTC 1213
DB |||||
DB 1 CGTGCATCTCACATGCACTACTACATGCTTGTTCACCGTTCGTCCTTGTTCATCGTC 60
QY |||||
QY 1214 CAAGCCTTGCCCTTCTTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 1273
DB |||||
DB 61 CAAGCCTTGCCCTTCTTCTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 120
QY |||||
QY 1274 GCAACTTCATGCAAAACAGCAATATGTTTCTGTAAC 1311
DB |||||
DB 121 GCAACTTCATGCAAAACAGCAATATGTTTCTGTAAC 158
XX
RESULT 7
AAL15210/c
ID AAL15210 standard; cdna; 883 BP.
XX
AC AAL15210;
XX
DT 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 7667.
DE Human breast cancer; cell marker; cytosstatic; 88.
XX
XX Homo sapiens.
XX

PN WO200151628-A2.
XX
XX 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 1378; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
XX
Query Match 5.4%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 3.4e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
XX
QY 419 TGTCCACTATTGGCACTACCCAGAACAGATTCTTAAATAACCAAGTAATCAATCA 478
DB |||||
DB 873 TGTNNANNANNACTTAAACCTTTAANANTNNNTNNTNNAANNAATAATTTAAANN 814
QY |||||
QY 479 CTCGAAGCTATCATGTATGTTTTAAAGAAACATCTATTTAAACCAAGTCTCTTAAA 538
DB |||||
DB 813 TANNAAATTTTNNAAAACTAATTTTANNNTANTNNTATTTTNTTAAAAAANNNAAAA 754
QY |||||
QY 539 AACAGCATATTTGCAAGAGACAAATATATGTACGTTTCAACATCTAAGACGACA 598
DB |||||
DB 753 TTAANNNTTNTTANTTANTTAAACCAAAATTTTNTTAAAAAATTTTNTTAAAAANT 694
QY |||||
QY 599 AATTATATCGAAGGTAAGCTATGAGCTTCAGATTTTCTTTTCATCTTGTATTTTG 658
DB |||||
DB 653 AATANTNTAAATTTTNTTAAATNAAAAAATTTTAAATTTTAAACAAATNTTTTTT 634
QY |||||
QY 659 TTATGTTTATATACATTTTCTCTCTTCAATAGAGTGATTTTCTCGATTTTATA 718
DB |||||
DB 633 TTTNTTNNATTAATAAANNTTTTAAATTAATAAANNTTTTNTTAAANATATTTTAA 574
QY |||||
QY 719 AAATGACTATAAGCATTTTATATATTAAGACGACGATGCTAGATTCGTTCAAAA 778
DB |||||
DB 573 AAAAAAANNTTTTNTTATANTTATTAATAAATTTATTTNTTNTTTCTTAAAAANA 514
QY |||||
QY 779 TCTTCTGATTTTNTTAAAGCTATGTTGGCAACCTGTTCTTCAAGATTTTGAT 838
DB |||||
DB 513 AAAAAAATTTTNTTAAANTTTTNTTAAACCTTAAATTTNAAANNTAATTTTNTTNT 454
QY |||||
QY 839 TTTTCAAAAAAATATGTTTATTTTCTTATTAATAAAGAAACATTAAGAAAAATACA 898
DB |||||
DB 453 NNANATTTAAAAATTTATTTTNTTANTTCTATTAATAAATTAANNTAATATATAT 394
QY |||||
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCATGCTGATATATT 958


```

Db      393 TTTNAAAAAATTAATTTATATNTATNTNAAANAATTTTAAATNTNANANAAAAATTTTAA 334
Qy      959 TGGCCAGCCCCCATTAATTTATTAACCGAAATCGAAATCGAGCAACCAATCTGAGCT 1018
Db      333 NTTAAATAAATAAATAATTAATTTAAACATCATTTNTAAATTTTAAATTAATAAAAAAAT 274
Qy      1019 ATTT 1022
Db      273 NTTT 270

RESULT 8
ACN85231/C
ID      ACN85231 standard; DNA; 960 BP.
XX
XX      ACN85231;
XX
XX      02-DEC-2004 (first entry)
XX
XX      Breast cancer related marker, seq id 6381.
XX
XX      Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX      Homo sapiens.
XX
XX      US2003099974-A1.
XX
XX      29-MAY-2003.
XX
XX      18-JUL-2002; 2002US-00198846.
XX
XX      18-JUL-2001; 2001US-0306220P.
XX
XX      (MILL-) MILLENNIUM PHARM INC.
XX
XX      Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX      WPI; 2003-787014/74.
XX
XX      Novel isolated polypeptide associated with breast cancer, useful for
XX      detecting presence of polypeptide in sample, as a marker for breast
XX      cancer.
XX
XX      Disclosure; SEQ ID NO 6381; 36bp; English.
XX
XX      The invention relates to an isolated polypeptide (I) associated with
XX      breast cancer which is encoded by a nucleic acid molecule comprising a
XX      nucleotide sequence (S1). Further disclosed is an antibody that binds to
XX      the polypeptide of the invention. The activity of the polypeptide of the
XX      invention may be described as cytostatic. The antibody is useful for
XX      detecting the presence of (I) in a sample. Nucleic acid molecules of the
XX      invention are useful in the detection of breast tumors. (I) is useful as
XX      a marker for breast cancer and in breast cancer therapy. Sequences given
XX      in records ACN78851-ACN92934 represent nucleic acid markers associated
XX      with breast cancer. Note: The sequence listing does not form part of the
XX      specification but may be obtained in electronic format from the USPTO web
XX      site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
XX      Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
XX
Query Match      5.4%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 3.5e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
Qy      419 TGTCCACTATGGCAGTACCGAGACAGATTTAAAAAATAACCAAGTAATATCCCA 478
Db      933 TGTNNNNNNANACCTTAACCTTTAAANATNTNNNATNTNANAAAAAATATTTAAANN 874
Qy      479 CTCGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACGATCCTTTAAAA 538
Db      873 TAAATAATTTTAAAAAATAATTTTAAANNAATNTNATATTTTNTTAAAAAANNNAAAAA 814

```

```

Qy      539 AACAGCATATTTTCGAAGAAGACAAATTAATGTTACAGTTTACAAACATCTAAGAGGACA 598
Db      813 TTAANNTTTTNTNANTTAATTAATAACCAAAATTTTTTAAAAAATTTTAAAAANTT 754
Qy      599 AATTATATCGAAAGTAGATGATGAGTTCAGATTTTCTTTTCATCTGTTATTTTG 658
Db      753 AATTAATTAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 694
Qy      659 TTATGTTTAAATTAATTAATTTCTCTTACAAATAGATGATTTTCTTCGATTTTAA 718
Db      693 TTTNTTTNNNAATAAAAAAATTTTAAATFAAAAAAANNANNTTTTATNANANATATTTAA 634
Qy      719 AATGACTATTAAGTATTTTATATAAGACGAGATGTCGAGATTCCTGCAAAA 778
Db      633 AAAAAAANNNTTTTTTTTAAATNTTAATAAAAAAATTTATTTNTTTTTCTNAAAAANA 574
Qy      779 TCTTCTGATTTTAAAGAGCTAGTTTGCAACCTGTTTCTTCAAGAATTTGATT 838
Db      573 AAAAAAATTTTNTNANNTTTTAAACCTTNAATTTNANANAAATTTTNNNTT 514
Qy      839 TTTTCAAAAAAATTAATTTTCTCTTATTAATAAAGAAACCTTAGAAAAATAGA 898
Db      513 NNNATTTAAAAAATTAATTTTNTCTATTAATTAATAAANNAATTAATAAT 454
Qy      899 GTTCCAGACTAGCCTTAGAATGTTTCCCAATTAATTAATCAATCACTGTATTAATT 958
Db      453 TTTNAAAAAATTAATTTATNTATNTNANANAAATTTAAANTNANANAAAAATTTTA 394
Qy      959 TGGCCAGCCCCCATTAATTTATTAACCGAAATCGAAATGAGGAGAAACCAATCTGAGCT 1018
Db      393 NTTTAAATTAATTAATTAATTTAAACATTTNTTAATTTTAAATTAATAAAAAAAT 354
Qy      1019 ATTT 1022
Db      333 NTTT 330

```

```

RESULT 9
ID      ADR04296 standard; DNA; 13400 BP.
XX
XX      ADR04296;
XX
XX      04-NOV-2004 (first entry)
XX
XX      Corn FT homologue nucleotide sequence SEQ ID NO:63.
XX
XX      flowering locus T gene; FT; terminal flower; TFL; Apetala3; Ap3; plant;
XX      floral development; plant sterility; plant fertility; flowering time;
XX      plant growth rate; inflorescence architecture; tissue culture morphology;
XX      cell division; FT homologue; gene; ds.
XX
XX      Zea mays.
XX
XX      PN MO2004067723-A2.
XX
XX      12-AUG-2004.
XX
XX      29-JAN-2004; 2004WO-US002422.
XX
XX      30-JAN-2003; 2003US-00343477.
XX
XX      (PION-) PIONEER HI-BRED INT INC.
XX      (DUPO ) DU PONT DE NEMOURS & CO B I.
XX
XX      Danilevskaya O, Hermon P, Bruggemann E, Shitroun D, Ananiev E;
XX      Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;
XX
XX      WPI; 2004-580996/56.
XX
XX      New polynucleotides, specifically nucleic acid fragments encoding
XX      flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
XX      homologs, useful for floral development, e.g. engineering plant flowering

```


PT time.
XX
PS Claim 6; SEQ ID NO 63; 109pp; English.
XX
CC The present invention describes an isolated polynucleotide comprising a
XX first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.
XX
SQ Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;
XX
XX
Query Match 5.3%; Score 69.8; DB 13; Length 13400;
Best Local Similarity 73.6%; Pred. No. 8e-05; Indels 0; Gaps 0;
Matches 89; Conservative 0; Mismatches 32;
QY 811 ACCGCTGTTCTTCAAGAAATTTGATTTTTCAAAAAATAGTTATTTCTCTTA 870
DB 8559 ACCATTTTATTTCAAGAGTTTATTTTATTCAGAAAAATAGTCAATTTCTCTTG 8618
QY 871 TAAATAGAAAAACATTAGAAAAATAGAGTTGCCAGACTAGCCCTTAGAATGTTTCCAA 930
DB 8619 AAAAATTAATAATCCATTAGAAAAATGGGTTGCAAACTAGTCTTATTAGTTTCCAT 8678
QY 931 T 931
DB 8679 T 8679
RESULT 10
AAZ10551
ID AAZ10551 standard; DNA; 2657 BP.
XX
AC AAZ10551;
XX
DT 16-NOV-1999 (first entry)
XX
DE DNA sequence of the P-Zeap promoter of maize.
XX
KW P-Zeap promoter; male sterile plant; glyphosate tolerance; glyphosate;
KW male reproductive tissue; hybrid seed production; crop outcrossing;
KW flower life; ss.
XX
XX Zea mays.
XX OS
XX FN WO9946396-A2.
XX PD 16-SEP-1999.
XX PF 09-MAR-1999; 99WO-US005126.
XX PR 09-MAR-1998; 98US-0072777P.
XX

PA (MONS) MONSANTO CO.
XX
XX PI Brown SM, Fromm ME;
XX
DR WPI; 1999-551420/46.
XX
PT Production of male sterile plants using a gene encoding glyphosate
XX tolerance, used for, e.g. production of hybrid seed.
XX
PS Disclosure; Fig 1A-B; 5app; English.
XX
XX The present sequence represents the P-Zeap promoter of maize. It is used
XX in the method of the invention. The specification describes a method for
XX the production of male sterile plants. The method comprises selective
XX expression of DNA encoding a protein that causes tolerance to glyphosate
XX and application of glyphosate. The method uses two DNA molecules, each
XX operably linked to a separate promoter, whereby the first promoter
XX functions in plant cells to produce a protein that causes tolerance to
XX glyphosate, and the second promoter functions in plant cells to cause the
XX production of a second RNA sequence in a male reproductive tissue.
XX Expression of the DNA promotes tolerance to glyphosate in those tissues
XX in which it is expressed. Expression of the second DNA molecule causes
XX the production of an RNA sequence which can inhibit the glyphosate
XX tolerance generated by expression of the first DNA molecule. By using a
XX promoter for the second DNA molecule which restricts the production of
XX the antisense RNA to only a subset of the tissues which express the first
XX DNA molecule, only the subset of tissues in which the second DNA molecule
XX is expressed will be susceptible to glyphosate toxicity. In this way, a
XX specific cell type or combination of cell types, depending upon the
XX promoters utilized, can be selectively ablated by application of
XX glyphosate to the plant. The methods can be used for producing male-
XX sterile plants for use in the production of hybrid seed, for minimizing
XX undesirable crop outcrossing, and for lengthening flower life. The
XX methods can be used with plants such as corn, wheat, rice, canola, oat,
XX barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
XX sunflower, soybean, tomato, cucumber and squash
XX
SQ Sequence 2657 BP; 666 A; 611 C; 677 G; 683 T; 0 U; 0 Other;
XX
XX
Query Match 5.1%; Score 67.2; DB 2; Length 2657;
Best Local Similarity 74.2%; Pred. No. 0.00019;
Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;
QY 787 ATTTTAAAGAGCAGAGTTTGCAACCTGTTCTTTCAGAAATTTGATTTTTCAAA 846
DB 1883 ATCTTTTAAGGCGCAGAGTTTGGAACACATTT-TTCCAAGGATTTCAATTTTGGCAG 1941
QY 847 AAAAATTAGTTATTTCTCTTTATTAATAAGAAAAACATTAGAAAAATAGAGTTGCCAG 906
DB 1942 GGAATTAGTTCATTTTCCCTTGGGAAATAGAAATCCATGGAAAAATGTGTTCCAA 2001
QY 907 ACTAGCCCTAGA 918
DB 2002 ACTAGCCCTAAA 2013
RESULT 11
AAK58751/C
ID AAK58751 standard; cDNA; 6027 BP.
XX
AC AAK58751;
XX
DT 16-AUG-1999 (first entry)
XX
DE Maize d111 gene encoding starch synthase enzyme D11.
XX
XX Starch synthase; SSII; D11 gene; maize; transgenic plant; ss.
XX
XX Zea mays.
XX OS
XX FN
XX PD
XX PF
XX PR
XX Key Location/Qualifiers
XX misc_feature 1..1437
XX FT /tag= f

FT	/note= "functional fragment of cDNA specifically claimed in Claim 16"
FT	in Claim 16"
FT	120. .5147
FT	/*tag= a
FT	120. .1221
FT	/*tag= b
FT	/note= "functional fragment of cdna specifically claimed in Claim 12"
FT	565. .816
FT	/*tag= d
FT	/note= "functional fragment of cdna specifically claimed in Claim 14"
FT	655. .1221
FT	/*tag= c
FT	/note= "functional fragment of cdna specifically claimed in Claim 15"
FT	1438. .2424
FT	/*tag= g
FT	/note= "functional fragment of cdna specifically claimed in Claim 17"
FT	2425. .3791
FT	/*tag= h
FT	/note= "functional fragment of cdna specifically claimed in Claim 18"
PN	
XX	N09924575-A1.
XX	
PD	20-MAY-1999.
XX	
PF	12-NOV-1998; 98MO-US024225.
XX	
PR	12-NOV-1997; 97US-00968542.
XX	
PA	(IOWA) UNIV IOWA STATE RES FOUND INC.
PI	
DR	Myers AM, James MG;
XX	
WP	WPI, 1999-327406/27.
DR	P-PDSB; AAY06199.
XX	
PT	
XX	
PS	Nucleic acid encoding starch synthase enzyme of maize.
ES	Claim 1, Page 104-107, 138pp; English.
XX	
CC	This is the nucleotide sequence of the maize gene dui1 (du1). To
CC	illustrate the role of the du1 locus in starch biosynthesis, a transposon
CC	tagging strategy was used to isolate the gene and describe its
CC	polypeptide product. The invention reports tagging of the du1 locus with
CC	Mu transposon, cloning and characterization of a portion of the gene, and
CC	isolation a near full-length cDNA (the present sequence). The amino acid
CC	sequence (see AAY06199) deduced from this cDNA indicates that Dui codes
CC	for a 186 kDa protein extremely similar to potato tuber starch synthase
CC	SSIII. Its expression pattern indicates that Dui codes for SSII of maize
CC	endosperm. The Dui product contains unique sequence features in its N-
CC	terminus that may mediate direct interactions with other starch
CC	biosynthetic enzymes. Mutations within the maize SSII gene affect
CC	multiple aspects of starch biosynthesis by disrupting an enzyme complex
CC	containing starch synthase(s), starch branching enzyme(s) and possibly an
CC	starch debranching enzyme(s). The isolated cDNA can be used to provide an
CC	enzyme with which to regulate the production of starch, and with which to
CC	produce altered or novel forms of starch, e.g. in transgenic plants.
CC	Expression of Dui in bacteria and yeasts also modifies glycogen
CC	production. Claimed expression vectors comprise the cDNA or fragments of
CC	it that code for functional portions of Dui
XX	
SQ	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
Query Match	5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity	77.5%; Pred No. 0.00035;

Matches	93;	Conservative	0;	Mismatches	26;	Indels	1;	Gaps	1;
Qy	792	TTTAAAGACTGAGTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTTTCACAAAAAAA	851						
Db	5606	TCCTACGGGCTTGTTGGGAACCCCATTT-TTCCAAAGGATTTTCATTTTTCGAAAGAAAA	5548						
Qy	852	TTAGTTATTTTCTCTTTATTAATAATAGAAAAACCTTAGAAAAATATGAGTTGCCAGACTAG	911						
Db	5547	TTAGTTATTTTTCATTTGAGAAAAATTAATAATCTCTTGAAAAAATATGAGTTACATACATCTAG	5488						
RESULT 12									
ID	ABX09935/c								
XX	ABX09935	standard; DNA; 6027 BP.							
XX	AC								
XX	ABX09935;								
DT	17-FEB-2003	(first entry)							
DE	DNA encoding	maize Starch synthase III (Dui).							
XX	KM	Starch; starch synthase; glucan association domain; GLASS; linker domain,							
XX	KM	LINKR; glucosyl transferase domain; GLYTR; C-terminal end; CTEND;							
XX	KM	granule bound starch synthase; GBSS; morphology; retrogradation;							
XX	OS	waterbinding; swelling potential; gene; ds.							
XX	OS	Zea mays.							
XX	WO200279410-A2.								
PD	10-OCT-2002.								
XX	29-MAR-2002;	2002WO-US009574.							
PF	30-MAR-2001;	2001US-0279720P.							
PR	(BADI) BASF PLANT SCI GMBH.								
PA	Communi P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;								
PI	WPI; 2003-040678/03.								
XX	New DNA encoding	fusion protein consisting of 4 different functional							
PT	domains selected	from glucan association domain, linker domain, glucosyl							
PT	transferase domain, and C-terminal end, useful for producing modified								
PT	starches.								
XX	Claim 33;	Page 225-227; 265pp; English.							
PS	The invention describes	an isolated DNA molecule encoding a fusion							
XX	protein consisting	of 4 different functional domains selected from glucan							
CC	association domain (GLASS), linker domain (LINKR), glucosyl transferase								
CC	domain (GLYTR), and C-terminal end (CTEND) which are operably linked to								
CC	one another. The DNA molecule is useful for expressing in plants								
CC	polypeptides including starch synthase enzymes as fusion proteins with								
CC	improved affinity to starch and modified catalytic capabilities and to								
CC	the in vivo and in vitro synthesis of glucan chains of modified lengths								
CC	as compared to plants producing native starch or starch produced with								
CC	native starch synthases. Expression of the starch synthase fusion								
CC	proteins along with granule bound starch synthase (GBSS) will lead to a								
CC	modified starch having an altered or improved morphology, retrogradation,								
CC	waterbinding, or swelling potential of the granules, gel strength,								
CC	adhesiveness, cohesiveness, hardness, elasticity, increased or decreased								
CC	granule size, degree of branching, crystallinity, degree of cross-								
CC	linking, and increased or decreased glucan chain lengths. This sequence								
CC	encodes a starch synthase used in the invention								
XX	Sequence 6027 BP, 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;								
Qy	Query Match	5.1%; Score 66.4; DB 8; Length 6027;							
XX	Best Local Similarity	77.5%; Pred. No. 0.00035;							
XX	Matches 93; Conservative	0; Mismatches 26; Indels 1; Gaps 1;							

QY	792	TTTAAAGACTGTTTGGGAACCCGTCTTTCCTTCAAAAGATTTTGATTTTTTCAAAAAA	851
QY	5666	TTCTAGGGGCTAGTTTGGGAACCCCAATT-TTCCAAAGGAGTTTCCATTTTCCAGAAAAA	5548
Db	852	TTAGTTTATTTTCTCTTATATAAATAGAAAAACCTTAGAAAAATAGAGTTGCCAGACTAG	911
QY	5547	TTAGTTTATTTTTCATTTGAAAAAATTTGAATCTCTTGGAAAAAATAGAGTTCCACTATACAG	5488
Db			

	RESULT 13
ID	ADK12106/c
ID	ADK12106 standard; cDNA, 6027 BP.
XX	
AC	ADK12106;
XX	
DT	03-JUN-2004 (first entry)
XX	
DE	cDNA encoding maize starch synthase, DUL.
XX	
KW	Maize; dul1 1; DUL; starch synthase;
KM	alpha-L,4-glycosyltransferase catalytic activity; starch production;
XX	plant; gene; ss.
OS	Zea mays.
XX	
FH	Key Location/Qualifiers
FT	CDS 120..514
FT	/tag= a
FT	/product= "DUL"
PV	US2004049810-A1.
XX	
PD	11-MAR-2004.
PF	05-AUG-2003; 2003US-00634262.
PR	12-NOV-1997; 9TUS-00968467.
PR	12-MAY-2000; 2000US-00554467.
XX	
PA	(MYER/) MYERS A M.
PA	(JAME/) JAMES M G.
XX	
PI	Myers AM, James MG;
PS	WI; 2004-238526/22.
DR	P-PSDB; ADK12117.
PT	New nucleic acid designated dul1, encoding a starch synthase, useful in starch production.
XX	
PS	Claim 1; SEQ ID NO 1, 58pp; English.
XX	
CC	The present invention relates to the isolation of a maize gene, dul1 1 (DUL), and the polypeptide it encodes. The DUL polypeptide has starch synthase activity, and comprises an N-terminal arm region, a C-terminal catalytic region, and a region of about 900 amino acids terminating with the catalytic region. The C-terminal catalytic region has a catalytic domain comprising alpha-L,4-glycosyltransferase catalytic activity. The dul1 polynucleotide sequence is useful in producing starch e.g. from a transgenic plant or transfected cell. The present sequence encodes maize DUL.
CC	
CC	
SO	Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;
Query Match	5.1%; Score 66.4; DB 12; Length 6027;
Bect Local Similarity	77.5%; Pred. No. 0.00035;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;	
OY	792 TTAAAGACTGATTGGCAACCGTTCTTCCTTCGAAGAATTTGTATTTTTCAAAAAA 851
Db	5606 TCTACGGGCTTGTTGGACAACCCTCATTT-TTCGAAGGATTTCCAATTTTTCGAAGAAA 5548
OY	852 TTACTTATTTTCTCTTNTATAAATAGAAAAACATTAGAAAAATAGATTGCCAGACTAG 911

Db 5547 TTAGTTTATTTTCATTGGAAAAATGTAAATCTCTGGAAAAAATGAGTTCACTACTAG 5488

RESULT 14
 ID ABX35844 standard; cDNA; 439 BP.
 XX ABX35844
 AC
 XX ABX35844;
 DT 20-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #1009.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 OS Bos Taurus.
 XX
 PN US2002137139-A1.
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 PA (BYATT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WAR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 PT New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 1009; 245pp; English.
 XX
 CC The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMPD), derived from
 CC cattle, and the LMPD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 1512 nucleotide sequences,
 CC appearing as ABX34836-ABX39947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of the molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any a
 CC of the 1512 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMPD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 1512 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=200201371139
 XX
 SO Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 198.304 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311

Sequence: 1 ccattggtcctctatcgaaaa.....cgcacatgcttctcgaaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	3	US-08-880-499-1
2	1311	100.0	1394	3	US-08-880-499-2
3	66.4	5.1	6027	2	US-08-968-542C-1
4	66.4	5.1	6027	4	US-09-554-467A-1
5	58.6	4.5	2523	2	US-08-410-784A-3
6	57	4.3	19124	2	US-08-487-826B-13
7	55.8	4.3	279	4	US-09-313-294A-5397
8	55.6	4.2	612	4	US-09-902-540-1357
9	55.6	4.2	1394	3	US-08-880-499-1
10	55.6	4.2	1394	3	US-08-880-499-2
11	55.4	4.2	1055	4	US-09-806-708B-23
12	55.2	4.2	2614	4	US-09-004-056-1
13	55	4.2	396	4	US-09-640-173-53
14	55	4.2	396	4	US-09-713-550-53
15	55	4.2	396	4	US-09-825-294-53
16	55	4.2	396	4	US-09-970-866-53
17	54.6	4.2	55886	4	US-09-949-016-15129
18	53.8	4.1	307	4	US-09-313-294A-4743
19	53.8	4.1	6027	2	US-08-968-542C-1
20	53.8	4.1	6027	4	US-09-554-467A-1
21	52.2	4.0	1141	4	US-09-806-708B-22
22	51.6	3.9	134987	4	US-09-949-016-15348
23	51.6	3.9	134987	4	US-09-949-016-15349
24	51.6	3.9	134987	4	US-09-949-016-15350
25	51.6	3.9	134987	4	US-09-949-016-15507
26	51.6	3.9	134987	4	US-09-949-016-15508
27	51.6	3.9	134987	4	US-09-949-016-15509

C	28	51.4	3.9	50383	4	US-09-949-016-17600	Sequence 17600, A
C	29	51.4	3.9	129415	4	US-09-949-016-16997	Sequence 16997, A
C	30	51	3.9	601	4	US-09-949-016-156535	Sequence 156535, A
C	31	51	3.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
C	32	50.6	3.9	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C	33	50.6	3.9	16573	4	US-09-949-016-14876	Sequence 14876, A
C	34	50.6	3.9	18773	4	US-09-949-016-14164	Sequence 14164, A
C	35	50.4	3.8	731	1	US-08-451-405A-2	Sequence 2, Appl1
C	36	50.2	3.8	832	4	US-09-621-976-2813	Sequence 2813, Ap
C	37	50.2	3.8	117366	4	US-09-949-016-16001	Sequence 16001, A
C	38	50	3.8	1392	3	US-09-257-584-1	Sequence 1, Appl1
C	39	49.8	3.8	2435	3	US-09-306-593-1	Sequence 1, Appl1
C	40	49.8	3.8	231129	4	US-09-949-016-16110	Sequence 16110, A
C	41	49.8	3.8	26293	4	US-09-949-016-11934	Sequence 11934, A
C	42	49.6	3.8	640681	4	US-09-790-988-1	Sequence 1, Appl1
C	43	49.2	3.8	126176	4	US-09-949-016-16137	Sequence 16137, A
C	44	49.2	3.8	126176	4	US-09-949-016-16138	Sequence 16138, A
C	45	47.8	3.6	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
; Sequence 1, Application US/08880499

; Patent No. 6037523

; GENERAL INFORMATION:

; APPLICANT: Albertson, Marc C.

; APPLICANT: Fox, Tim W.

; APPLICANT: Carl, Garnaat W.

; APPLICANT: Hultman, Gary A.

; APPLICANT: Kendall, Timmy L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

; CITY: Johnston

; STATE: Iowa

; COUNTRY: USA

; ZIP: 50131

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: CONCURRENTLY HEREWITH

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Sweeney, Patricia A.

; REGISTRATION NUMBER: 32,733

; TELEPHONE/DOCKET NUMBER: 0578

; TELEPHONE: (515) 248-4800

; TELEFAX: (515) 248-4844

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1394 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-880-499-1

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1,1e-313;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGAGTCTCTATGAAAAAGATGATCAATGTCATATCCGTTTCTTAAAGGTC 60
Db 1 CCATGAGTCTCTATGAAAAAGATGATCAATGTCATATCCGTTTCTTAAAGGTC 60
QY 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
QY 121 CTCATGTTTCACTTCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACACTTTGCGGACACCCATCAAGGCTTTGCGATGCGCA 240
Db 181 CTGACCGAGGCCCATGACACACTTTGCGGACACCCATCAAGGCTTTGCGATGCGCA 240
QY 241 CGAGAGCTATCGGGTGTGATCCAGGGGATATATGTCCTCCCAATGCTACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGATCCAGGGGATATATGTCCTCCCAATGCTACCTATA 300
QY 301 TTATTAATCTTATTAATTAATTTTGAATAAACAATTAATCTTATCTTTGTA 360
Db 301 TTATTAATCTTATTAATTAATTTTGAATAAACAATTAATCTTATCTTTGTA 360
QY 361 GGGCTCAGCAATGATTTTCTGTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
Db 361 GGGCTCAGCAATGATTTTCTGTTAGGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
QY 421 TCCACTATTTGGCATCCAGAACAGATTTTAAAAAATACCAAGTAACTATCACT 480
Db 421 TCCACTATTTGGCATCCAGAACAGATTTTAAAAAATACCAAGTAACTATCACT 480
QY 481 CGAAGCTATCATGATGATTTTAAAGAAACATCTATTAACACGATCCTTAAAAAA 540
Db 481 CGAAGCTATCATGATGATTTTAAAGAAACATCTATTAACACGATCCTTAAAAAA 540
QY 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAA 600
Db 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGCGACAA 600
QY 601 TTATATCGAAAGTAGTATGAGTCGATGATTTTCTTTCATCTTGTATTTTGT 660
Db 601 TTATATCGAAAGTAGTATGAGTCGATGATTTTCTTTCATCTTGTATTTTGT 660
QY 661 ATTGTTTATATACATTTTCTCTCTCAATAGATGATTTTCTCGATTTTAA 720
Db 661 ATTGTTTATATACATTTTCTCTCTCAATAGATGATTTTCTCGATTTTAA 720
QY 721 ATGACTATAAAGTCAATTTTATATAAGACGCGATGCGTAGATTCGTTCAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATAAGACGCGATGCGTAGATTCGTTCAAAATC 780
QY 781 TTTTGAATTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTTGAATTTTAAAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
QY 841 TTTCAAAAAAATAGTTTATTTTCTTATTAATAATAGAAAACTTAGAAAAATAGAT 900
Db 841 TTTCAAAAAAATAGTTTATTTTCTTATTAATAATAGAAAACTTAGAAAAATAGAT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTGATATTTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTGATATTTTG 960
QY 961 GCGAGCCCATTAATTAATTAACGGAATCTGAATCGGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGCCCATTAATTAATTAACGGAATCTGAATCGGAAACCAATCTGAGCTAT 1020
QY 1021 TTTCTAGATTAGTAAAAAGGAGAGAGAGAAATCAATTGATTAAGTGTGCC 1080
Db 1021 TTTCTAGATTAGTAAAAAGGAGAGAGAGAAATCAATTGATTAAGTGTGCC 1080
QY 1081 TGAATGTGCGGTTTGGCAACGATAGCACCGTAAATCATAGTCAATAGTGTGCTACGTCA 1140

Db 1081 TGAATGTGCGGTTTGGCAACGATAGCACCGTAAATCATAGTCAATAGTGTGCTACGTCA 1140
QY 1141 GGTGGGCGAGCTCTGTGTCATCTCAATGAGATACATGCTTTGTTACACCGTTGTC 1200
Db 1141 GGTGGGCGAGCTCTGTGTCATCTCAATGAGATACATGCTTTGTTACACCGTTGTC 1200
QY 1201 TTGTTTCATGCTGCAAGCCTTCTATTCTGAACCAAGAGATACCTACTCCCAACAT 1260
Db 1201 TTGTTTCATGCTGCAAGCCTTCTATTCTGAACCAAGAGATACCTACTCCCAACAT 1260
QY 1261 CCATCTTAATCTGCAATCTTCCATGCAACGACGACATATGTTTCTGAC 1311
Db 1261 CCATCTTAATCTGCAATCTTCCATGCAACGACGACATATGTTTCTGAC 1311

RESULT 2
US-08-880-499-2
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garneat W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Jimmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OR INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 100.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1,1e-313;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGAGTCTCTATGAAAAAGATGATCAATGTCATATCCGTTTCTTAAAGGTC 60
Db 1 CCATGAGTCTCTATGAAAAAGATGATCAATGTCATATCCGTTTCTTAAAGGTC 60
QY 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
Db 61 CTTCTTCTGCTTATTAATGATCGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120


```

QY 121 CTCGATGTTCCACTTCTCCCACTCGGTTGCACTTCTTGATGTCGGTTCCTAT 180
DB 121 CTCGATGTTCCACTTCTCCCACTCGGTTGCACTTCTTGATGTCGGTTCCTAT 180
QY 181 CTGACCGAGGCCCATGAGACCTTTGCGGACACCACTCAAGGGCTTTGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGAGACCTTTGCGGACACCACTCAAGGGCTTTGATGGCCCA 240
QY 241 CGAGACGTATCGGCTCGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
DB 241 CGAGACGTATCGGCTCGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
QY 301 TTATTTATCTTTAGATATATATTTTGGAAAAATACAACTTATCTTTGTGTA 360
DB 301 TTATTTATCTTTAGATATATATTTTGGAAAAATACAACTTATCTTTGTGTA 360
QY 361 GGGCTCAGCATAGATTTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
DB 361 GGGCTCAGCATAGATTTTTCGTTAGGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAAATTTAAAAATTAACCAAGTAACTATCCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAAATTTAAAAATTAACCAAGTAACTATCCACT 480
QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCTCTTAAAAA 540
DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCTCTTAAAAA 540
QY 541 CAAGCATATTTGCAAAAGAGACAAATATGTTACAGTTTCAAAATCTTAAGAGCGAACA 600
DB 541 CAAGCATATTTGCAAAAGAGACAAATATGTTACAGTTTCAAAATCTTAAGAGCGAACA 600
QY 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTTATTTGTT 660
DB 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTTATTTGTT 660
QY 661 ATTGTTTTATATACATTTTCTTCTTACATATAGTGATTTTCTTCGATTTTAAAA 720
DB 661 ATTGTTTTATATACATTTTCTTCTTACATATAGTGATTTTCTTCGATTTTAAAA 720
QY 721 ATGACTATTAAGTCAATTTTATATTAAGAGACGATGCTAGATCTCGTTCAAAAAATC 780
DB 721 ATGACTATTAAGTCAATTTTATATTAAGAGACGATGCTAGATCTCGTTCAAAAAATC 780
QY 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTCTTTTCAAGAATTTGATTTT 840
DB 781 TTTCTGATTTTAAAGAGCTAGTTTGGCAACCTGTCTTTTCAAGAATTTGATTTT 840
QY 841 TTCAAAAAAAATTAAGTTATTTTCTTCTTATTAAGAAACACCTTAAAGAAATAGAGT 900
DB 841 TTCAAAAAAAATTAAGTTATTTTCTTCTTATTAAGAAACACCTTAAAGAAATAGAGT 900
QY 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGATTAATTTTG 960
DB 901 TGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGATTAATTTTG 960
QY 961 GCCAGCCCCATAAATTAATTTAAACGAAACTGAATCGAGCGAAACCAATCTGACTAT 1020
DB 961 GCCAGCCCCATAAATTAATTTAAACGAAACTGAATCGAGCGAAACCAATCTGACTAT 1020
QY 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGTCACTGTCCC 1080
DB 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCACTTTTAAAGTCACTGTCCC 1080
QY 1081 TGAGATGTGCGGTTGGCAACGATAGCACCGTATCATAGCTATAGGCTAGCTAGTCA 1140
DB 1081 TGAGATGTGCGGTTGGCAACGATAGCACCGTATCATAGCTATAGGCTAGCTAGTCA 1140
QY 1141 GGTTCGGACGCTCTCGTGCATCTCAATGCGCATCACTAGCTTGTTCACGCTTCGTC 1200
DB 1141 GGTTCGGACGCTCTCGTGCATCTCAATGCGCATCACTAGCTTGTTCACGCTTCGTC 1200

```

```

QY 1201 TTGTTCCATGTCGCAAGCCTTGCTATCTGTAACCAAGAGATTACTACTCCCAACAAT 1260
DB 1201 TTGTTCCATGTCGCAAGCCTTGCTATCTGTAACCAAGAGATTACTACTCCCAACAAT 1260
QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGTAAC 1311

RESULT 3
US-08-968-542C-1/c
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dU11 Codes for A No. 5981728el Starch
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHEICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm
; IMMEDIATE SOURCE:
; LIBRARY: maize endosperm cDNA library in
; LIBRARY: gtl1
; CLONE: pmg10; pmg6aa; pmg6-2M
; US-08-968-542C-1

Query Match 5.1%; Score 66.4; DB 2; Length 6027;
Best Local Similarity 77.5%; Pred. No. 2.4e-06;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
QY 792 TTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
DB 792 TTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
QY 5606 TCTACGGGCTAGTTGGAAACCCCAATTT-TTCAAGGGATTTTCAATTTTCCAAAGAAAA 5548
DB 5606 TCTACGGGCTAGTTGGAAACCCCAATTT-TTCAAGGGATTTTCAATTTTCCAAAGAAAA 5548
QY 852 TTAGTTTATTTCTCTTATATAAATGAAAAACCTTGAAGAAATGAGTTGCAAGCTAG 911
DB 852 TTAGTTTATTTCTCTTATATAAATGAAAAACCTTGAAGAAATGAGTTGCAAGCTAG 911

```


Db 5547 TTAGTTTATTTTCATTGGAAAAATTTGAATCTCTTGAAAAATAGAGTTACATACACTAG 5488

RESULT 4
US-09-554-467A-1/c
Sequence 1, Application US/09554467A
Patent No. 6639125
GENERAL INFORMATION:
APPLICANT: Myers, Alan M.
TITLE OF INVENTION: dulli Coding for a No. 6639125el Starch Synthase and Uses
FILE REFERENCE: D6036PCT
CURRENT APPLICATION NUMBER: US/09/554,467A
PRIOR FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: PCT/US98/24225
PRIOR FILING DATE: 1998-11-12
PRIOR APPLICATION NUMBER: US 08/062,102
NUMBER OF SEQ ID NOS: 37
SEQ ID NO 1
LENGTH: 6027
TYPE: DNA
ORGANISM: maize
FEATURE:
OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the
OTHER INFORMATION: starch synthase enzyme DUL.
US-09-554-467A-1

Query Match
Best Local Similarity 5.1%; Score 66.4; DB 4; Length 6027;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

Qy 792 TTAAAGCTAGTTCGCAACCCCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
Db 5606 TCTACGGCTAGTTTGGAACCCCATTT-TTCCAGAGATTTCATTTTCCAGAAAA 5548

Qy 852 TTAGTTTATTTCTCTTAAATAGAAAACTAGAAAAATAGATTGCGACACTAG 911
Db 5547 TTAGTTTATTTTCATTGGAAAAATTTGAATCTCTTGAAAAATAGAGTTACATACACTAG 5488

RESULT 5
US-08-410-784A-3/c
Sequence 3, Application US/08410784A
Patent No. 5912413
GENERAL INFORMATION:
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G.
TITLE OF INVENTION: ISOLATION OF SUL, A STARCH DEBRANCHING
TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PasteSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Heine, Ph.D., Holliday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2523 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-410-784A-3

Query Match
Best Local Similarity 4.5%; Score 58.6; DB 2; Length 2523;
Matches 105; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

Qy 774 AAAATCTTCTGATTTTAAAGAGCTAGTTGGCAACCCGTGTTCTTCAAGAAATT 833
Db 279 ACATATACATATACATATATAGGTCTAGTTGCAATCCATTTTATCAAAAG--TTT 222

Qy 834 TGATTTTTCAAAAAAATAGTTATTTTCTCTTATPAATAGAAAACTAGAA 893
Db 221 TACATTTTTCAAAAATAGTTATTTTCTCTTGA-AAATAGAAATTTCTCAGAAAA 163

Qy 894 ATAGAGTTCAGAGACTAGCCCTAGAT 920
Db 162 ATAGAGTTTACAACTAGCTTAAAT 136

RESULT 6
US-08-487-826B-13/c
Sequence 13, Application US/08487826B
Patent No. 5993827
GENERAL INFORMATION:
APPLICANT: Sim, Kim L.
APPLICANT: Chltnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
TITLE OF INVENTION: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 4.3%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00078; Indels 0; Gaps 0;

Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAAGCTATGACCTGAGATTTCTTTTCATCTGTATTTGTTATGTTTATAT 673

DB 15677 TAATGTTTTTTTTCTTCTTTGTTTTTATTTTATATCATTTTTTTTATAT 15618

QY 674 ACATTTCTCTCTTCAATAGAGTATTTCTCCGATTTATATAAGACTATAAGT 733

DB 15617 AAAATTTTTTTTAAATTTTTTTTGAAATCTTTTCATTTTATCTATCAAAATTTA 15558

QY 734 CATTTATATAGAGACGATGCGTATCTGTCATAAAATCTTTCTGATTTTTT 793

DB 15557 TATTTATATATATTTTATTTTATTTTAAATAATTTCTCTTTTTTTTTTTTT 15498

QY 794 TAAGACTAGTTGGACACCTGTTCTTCAAGATTTGATTTTTTCAAAAAAAT 853

DB 15497 TTTTATTTAAATTAATTTTTTTTATATTTTCATTTTCTTTTTCATTTTAAATTA 15438

QY 854 AGTTTATTTCTCTTTATATAAATAGAAAAACATTAGAAAA 894

DB 15437 GTTTTATATTTCTTTTAAATTAATTAATCATATATATAA 15397

RESULT 7

US-09-313-294A-5397

Sequence 5397, Application US/09313294A

Patent No. 6476212

GENERAL INFORMATION:

APPLICANT: Lalsudi, Raghunath V.

APPLICANT: Ito, Laura Y.

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

FILE REFERENCE: PL-0017 US

CURRENT APPLICATION NUMBER: US/09/313,294A

CURRENT FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 7600

SOFTWARE: PERL Program

SEQ ID NO 5397

LENGTH: 279

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID No. 6476212 700350078H1

NAME/KEY: unsure

LOCATION: 10, 12, 185-186, 204, 253, 274, 278

OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-5397

Query Match

Best Local Similarity 4.3%; Score 55.8; DB 4; Length 279;

Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 789 TTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAAATTTGATTTTTCAAAA 848

DB 3 TTCCTAANCCCTAGTTGGAAACCCCATTTTCCCAAGGTTTTCATTTTCCCAAGG 62

QY 849 AATATGTTATTTCTCTTTAT-AAATAGAAAAACATTAGAAAAATAGATGCCAGA 907

DB 63 AAGTAGAACATTTTCCCTTGGAATAAATCCTTGGGAAATGAGTTCCCAA 122

QY 908 CTAGCCCTA 916

DB 123 CTAGCCCTA 131

RESULT 8

US-09-902-540-1357/c

Sequence 1357, Application US/09902540

Patent No. 683447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

APPLICANT: Hinkle, Gregory J.

APPLICANT: Slater, Steven C.

APPLICANT: Miegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 1357

LENGTH: 612

TYPE: DNA

ORGANISM: Myxococcus xanthus

FEATURE:

NAME/KEY: unsure

LOCATION: (1)-(612)

OTHER INFORMATION: unsure at all n locations

US-09-902-540-1357

Query Match

Best Local Similarity 4.2%; Score 55.6; DB 4; Length 612;

Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCACTCTGTATTTGTTATTTGTTATATCATTTCTCT 685

DB 560 TTTTAAATTT 501

QY 686 CTACAAATGAGATTTCTCCGATTTTAAATGACTAAAGTATTTATATA 745

DB 500 TTTTATTT 441

QY 746 AGAGACGATGCTAGATTTCTCGTCAAAATCTTCTGATTTTAAAGCTAGT 805

DB 440 ATTTTATTTAT 381

QY 806 TGGCAACCTGTTCTTCAAGAAATTTGATTTTTCAAAAAATAGTTATTTCT 865

DB 380 TTTTATTTAN 321

QY 866 CTTTATA 872

DB 320 TATTTTA 314

RESULT 9

US-08-880-499-1/c

Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Hufman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match 4.2%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTGGGAAACATTCGAGGCTAGTGGCACTATTTT 891
QY 821 TTTCAAAGATTTGATTTTTCAAAAAATTAAGTTATTTCTTTATATAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAAATTAATAATTTTGGAAAAATCCAAA 831
QY 881 AACACTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTGGCAACCTAGCTTTAAAAAATCAGAAAGATTTTGA 771

QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761

RESULT 10
US-08-880-499-2/c
Sequence 2, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Albertson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garneal W.
APPLICANT: Huffman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match 4.2%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.00064;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGCTAGTTGGCAACCTGTTTC 820
DB 950 TACACAGTATGTATTTATTGGGAAACATTCGAGGCTAGTGGCACTATTTT 891
QY 821 TTTCAAAGATTTGATTTTTCAAAAAATTAAGTTATTTCTTTATATAATAGAA 880
DB 890 TCTAAGTGTCTTATTTTATAAGAAATTAATAATTTTGGAAAAATCCAAA 831
QY 881 AACACTTAGAAAAATAGAGTGGCAGACTAGCCCTAGAAATGTTTCCCAATTAATTACA 940
DB 830 TTCTTTGAAGAAACAGGGTGGCAACCTAGCTTTAAAAAATCAGAAAGATTTTGA 771
QY 941 TCACGTGTGA 950
DB 770 CGAGATCTA 761

RESULT 11
US-09-806-708B-23
Sequence 23, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent version 3.0
SEQ ID NO 23
LENGTH: 1055
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: Promoter
LOCATION: (1) .. (1055)
OTHER INFORMATION: consensus sequence of A.t. and L.a. FABI promoters
US-09-806-708B-23

Query Match 4.2%; Score 55.4; DB 4; Length 1055;
Best Local Similarity 22.5%; Pred. No. 0.00065;
Matches 182; Conservative 181; Mismatches 430; Indels 15; Gaps 3;

Oy	692	AATGAGGATTTCTTCGCCATTATTAATAAAGACTATTAAGCA	751
Db	71	TTT	130
Oy	752	CGCATGTGATCTCGTCCAAAACTTTCTGATTTTTTAAGACGATTTGGCAA	811
Db	131	TTT	190
Oy	812	CCCGTCTCTTCGAAGAATTGATTTTTCAGAAAAAAATTAGTTATTTCTCTAT	871
Db	191	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTANNNTTTTNNTTTNCCTTNTTTT	250
Oy	872	AAAATAGAAAACCTTAGAAAAATAGAGTTGCAGACTAGCCCTAGAAATGTTTTCCCAAT	931
Db	251	AATTCAPAAAAAANAANAGAAAANATPAANANNACNNANNNNNNATNTNCTTATA	310
Oy	932	AAATTCACATCACTGNG	948
Db	311	MTNNNTNNNNNANGG	327

```

RESULT 14
US-09-713-550-53
Sequence 53, Application US/09713550
Patent No. 6617109
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Seolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT OF OVARIAN CANCER
FILE REFERENCE: 210121.484C4
CURRENT APPLICATION NUMBER: US/09/713,550
CURRENT FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 205
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(396)
OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

```

Query Match	4.2%	Score 55	DB 4	Length 396
Best Local Similarity	43.8%	Pred. NO. 0.00056		
Matches 139, Conservative		0, Mismatches 178	Indels 0	Gaps 0

Qy	632	TTTTTCTTTTCATCTCTGTAAATTTGGTAATGTTTATATACATTTCTCCCTTACA	691
Db	11	TT	70
Qy	692	ATAGAGTGATTTCTTCCGATTTTATATAAAAGACATATAAGTCAATTTTATATAAGCA	751
Db	71	TT	130
Qy	752	CGCATGCGTAGATCTCGTCAAAAATCTTCTGATTTTTTTAAGACGTAGTTGGCA	811
Db	131	TT	190
Qy	812	CCCTGTTCTTTCATAAAGAAATTTGATTTTTTCAAAAAAATTTAGTTATTTCTCTTAT	871
Db	191	TT	250
Qy	872	AAATATGAAAAACCTTAGAAAAATAGAGTGGCAGACATAGCCTAGAATGTTTCCCAAT	931
Db	251	AATTCANAAAAAAGAAAAAANNTTAANANNANNNANNNNNNNNATNTNTCTTATA	310
Qy	932	AAATTACATCACTGTG	948
Db	311	NTNNNTNNNNNNNANNGG	327

```

RESULT 15
US-09-825-294-53
Sequence 53, Application US/09825294
Patent No. 6710170
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolk, John A.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
FILE REFERENCE: 210121,48405
CURRENT APPLICATION NUMBER: US/09/825,294
CURRENT FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 215
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(396)
OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

```

Query Match	4.2%	Score 55;	DB 4;	Length 356;
Best Local Similarity	43.8%;	Pred. No. 0.00056;		
Matches 139;	Conservative	0;	Mismatches 178;	Indels 0;
			Gaps	0

Oy	63	TTTTTCCTTTAACTCTGTATATTTGGTATGTTTTATATATACATTTCTCTCTACA	691
Db	11	TT	70
Oy	692	ATAGAGTATTTCTCCGATTTTATATAAATGACTATAAAGCATTTTATATAGACA	751
Db	71	TT	130
Oy	752	CGCATGTCGAGATTCCTGGTCAAAATCTTCTGATTTTTTTATAGAGCTAGTTGGCA	811
Db	131	TT	150
Oy	812	CCCTGTTCTTCAAGAATTTGATTTTTTCAAAAAATTAGTTATTTCTCTTAT	871
Db	191	TT	250
Oy	872	AAATATGAAGACACTTAGAAAAATAGAGTTGGCAGACTAGCCTAGAATGTTTTCCCAAT	931
Db	251	AATTCANAAAAAGAAAATAAANANNANNNANNNNNNATATNTTCTTATATA	310
Oy	932	AAATTCAAATCACTGTG	948
Db	311	NTNNTTNNNNNANNGG	327

Search completed: September 15, 2005, 08:25:15
Job time : 201.304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 853.804 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311
Sequence: 1 cccatgctgctctcatgaaa.....cgcacatagctctccgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 33128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09C_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 22: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 23: /cgn2_6/prodata/2/pubpna/US11_PUBCOMB.seq:*
- 24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1311	100.0	1394	20	US-10-713-381-1
2	1311	100.0	1394	20	US-10-713-381-2
3	160.6	12.3	255	20	US-10-713-381-9
4	146	11.1	158	20	US-10-713-381-3
5	80.8	6.2	1261	20	US-10-425-115-134230
6	74	5.6	320	18	US-10-425-114-23340
7	74	5.6	624	18	US-10-425-114-16264

8	73	5.6	1326	20	US-10-425-115-141826	Sequence 141826, A
9	72.8	5.6	2445	18	US-10-425-114-32493	Sequence 32493, A
10	72.8	5.6	2729	20	US-10-425-115-83293	Sequence 83293, A
11	72.4	5.5	1203	20	US-10-425-115-51470	Sequence 51470, A
12	71.8	5.5	1123	20	US-10-425-114-17816	Sequence 17816, A
13	71.8	5.5	1123	20	US-10-425-115-142853	Sequence 142853, A
14	71.8	5.5	2537	18	US-10-425-114-31957	Sequence 31957, A
15	71.8	5.5	2537	20	US-10-425-115-52216	Sequence 52216, A
16	71.8	5.5	3607	18	US-10-425-114-31061	Sequence 31061, A
17	71.8	5.5	3691	20	US-10-425-115-52219	Sequence 52219, A
18	70.2	5.4	960	14	US-10-198-846-6381	Sequence 6381, A
19	69.8	5.3	13400	21	US-10-343-477A-63	Sequence 63, App1
20	69.6	5.3	724	20	US-10-425-115-87756	Sequence 87756, A
21	69.2	5.3	527	18	US-10-425-114-4041	Sequence 4041, A
22	68.8	5.2	610	20	US-10-425-115-47165	Sequence 47165, A
23	68.4	5.2	1215	18	US-10-425-114-24656	Sequence 24656, A
24	68.4	5.2	2863	20	US-10-425-115-75310	Sequence 75310, A
25	67.8	5.2	512	20	US-10-425-115-49781	Sequence 49781, A
26	67.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
27	67.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
28	67.4	5.1	1260	18	US-10-425-114-30881	Sequence 30881, A
29	67.4	5.1	1431	20	US-10-425-115-149304	Sequence 149304, A
30	67.2	5.1	2232	18	US-10-425-114-16607	Sequence 16607, A
31	67.2	5.1	2249	18	US-10-425-114-20264	Sequence 20264, A
32	66.4	5.1	6027	18	US-10-634-262-1	Sequence 1, App1
33	66.4	5.1	6027	19	US-10-109-048-1145	Sequence 1145, App
34	66.4	5.1	6051	20	US-10-425-115-41589	Sequence 41589, A
35	66.2	5.0	439	9	US-09-960-352-1009	Sequence 1009, App
36	66	5.0	1376	18	US-10-425-114-2073	Sequence 2073, App
37	65.6	5.0	721	20	US-10-425-115-177955	Sequence 177955, App
38	65.6	5.0	2274	18	US-10-425-114-830	Sequence 830, App
39	65.6	5.0	2632	20	US-10-425-114-162957	Sequence 162957, App
40	65.6	5.0	8056	20	US-10-473-126-386	Sequence 386, App
41	65.6	5.0	1364	20	US-10-425-115-150820	Sequence 150820, A
42	65	5.0	833	20	US-10-425-115-30824	Sequence 30824, A
43	64.8	4.9	1838	20	US-10-425-115-21012	Sequence 21012, A
44	64.8	4.9	1838	20	US-10-425-115-48324	Sequence 48324, A
45	64.2	4.9	581	20	US-10-425-115-48324	Sequence 48324, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; TITLE OF INVENTION: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTCATGAAAGATGATGATGATGCTCTATATCCGTTTCTTAGGCTCC 60


```
Db 1 CCAATGATGCTCTATGAAAAAGATGATCAATGTCATATCCGTTTTCTTAAGGTC 60
QY 61 CTTCTTCTGCTTATTAAGTACATGATCGGGTTTACAAAACTTCCAGGGTCATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTACATGATCGGGTTTACAAAACTTCCAGGGTCATGAT 120
QY 121 CTCATGTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGCGGTGGTCCCAT 180
Db 121 CTCATGTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCCCTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCCCTTTCGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGCTACCTTAA 300
Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGCTACCTTAA 300
QY 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTGTGTA 360
Db 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTGTGTA 360
QY 361 GGGCTTCAGCATAGATTTTTCGCTTAAGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
Db 361 GGGCTTCAGCATAGATTTTTCGCTTAAGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
QY 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAAAAATACAAAGTAACTATCACT 480
Db 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAAAAATACAAAGTAACTATCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTTAAAAA 540
QY 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
QY 601 TTATATCGAAAGTAAGCTATGACGTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
Db 601 TTATATCGAAAGTAAGCTATGACGTTCAATTTTCTTTTCAATCTTGTATTTTGT 660
QY 661 ATGTTTTTATATACATTTTCTCTCAATAGAGTATTTCTTCCGATTTTAA 720
Db 661 ATGTTTTTATATACATTTTCTCTCAATAGAGTATTTCTTCCGATTTTAA 720
QY 721 ATGACTATAAAGTCATTTTATATAAAGACGCGATGCTAGATTTCTGTTCAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATAAAGACGCGATGCTAGATTTCTGTTCAAAATC 780
QY 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
Db 781 TTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTCAAGAAATTTGATTT 840
QY 841 TTCAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACACTTAGAAAAATAGAT 900
Db 841 TTCAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACACTTAGAAAAATAGAT 900
QY 901 TGCAGACTAGCCCTAGAAATGTTTTCCAAATTAATCAATCACTGTATATTTTG 960
Db 901 TGCAGACTAGCCCTAGAAATGTTTTCCAAATTAATCAATCACTGTATATTTTG 960
QY 961 GCGAGCCCATTAATTTTAAACCGAAACTGAAATCGAGCGAAACCAATCTGACCTAT 1020
Db 961 GCGAGCCCATTAATTTTAAACCGAAACTGAAATCGAGCGAAACCAATCTGACCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAATCAGTTTAAAGTCACTTGTCC 1080
Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAATCAGTTTAAAGTCACTTGTCC 1080
QY 1081 TGAGATGTGCGGTTTGGCAACGATGACCACTGATCATAGTGCCTACGTCA 1140
Db 1081 TGAGATGTGCGGTTTGGCAACGATGACCACTGATCATAGTGCCTACGTCA 1140
```

```
Db 1081 TGAGATGTGCGGTTTGGCAACGATGACCACTGATCATAGTGCCTACGTCA 1140
QY 1141 GGTTCGAGCTCTGTGTGATCTCAATGGAATCTATACCTGTTCAACCGTTGTC 1200
Db 1141 GGTTCGAGCTCTGTGTGATCTCAATGGAATCTATACCTGTTCAACCGTTGTC 1200
QY 1201 TTGTCATGCTCAAGCCTTGCTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Db 1201 TTGTCATGCTCAAGCCTTGCTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
QY 1261 CCATCTACTCATGCACTTCCATGCAACGACGACATATGTTTCTGAAC 1311
Db 1261 CCATCTACTCATGCACTTCCATGCAACGACGACATATGTTTCTGAAC 1311

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE--PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5e-275;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCAATGATGCTCTATGAAAAAGATGATCAATGTCATATCCGTTTTCTTAAGGTC 60
Db 1 CCAATGATGCTCTATGAAAAAGATGATCAATGTCATATCCGTTTTCTTAAGGTC 60
QY 61 CTTCTTCTGCTTATTAAGTACATGATCGGGTTTACAAAACTTCCAGGGTCATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTACATGATCGGGTTTACAAAACTTCCAGGGTCATGAT 120
QY 121 CTCATGTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGCGGTGGTCCCAT 180
Db 121 CTCATGTTCCACTTCCCACTCGGGTGGACATTTCTTGGATGCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCCCTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACACCTTTCGGGACACCATCAAGGCCCTTTCGATGGCCCA 240
QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGCTACCTTAA 300
Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGATATATGTCCTCCCAATGCTACCTTAA 300
QY 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTGTGTA 360
Db 301 TTATTAATCTTGAATATTAATTTTGGAAAAATACAACTTAATCTTGTGTA 360
QY 361 GGGCTTCAGCATAGATTTTTCGCTTAAGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
Db 361 GGGCTTCAGCATAGATTTTTCGCTTAAGGCCAGAAAATGCGAGGACCGCATGCTAGTG 420
QY 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAAAAATACAAAGTAACTATCACT 480
Db 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAAAAATACAAAGTAACTATCACT 480
```



```
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230
```

```
Query Match          6.2%; Score 80.8; DB 20; Length 1261;
Best Local Similarity 78.2%; Pred. No. 3.1e-07;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
Qy 792 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 851
    |||||
Db 999 TCTTAAGGCTAGTTGGAACTTATTTTAAAGATTTTCTTTTAAAGTAA 1058
Qy 852 TTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAATTAAGTTGCCAGACTAG 911
    |||||
Db 1059 TTAGTTATTTCTCTTGAAGAAATGAAATCTCTTGAAATTAAGTTGCTAAACTAG 1118
Qy 912 CCCT 915
    |||||
Db 1119 CCCT 1122
```

```
RESULT 6
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340
```

```
Query Match          5.6%; Score 74; DB 18; Length 320;
Best Local Similarity 73.1%; Pred. No. 4.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
Db 62 TTTTGGGCTAGTTGGAATCTCATTTTTCAGAGATTTTATTTCCAAAGAA 121
Qy 851 ATTAGTTATTTCTTATTAATAATGAAAAACATTGAAATTAAGTTGCCAGACTA 910
    |||||
```

```
Db 122 ATTAGTTATTTCTTGAAGAAATAGAAATCCCTTGGAATTAAGTTCTAAACGA 181
Qy 911 GCCCTAAGT 920
    |||||
Db 182 GCCCTAATT 191
```

```
RESULT 7
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E.
/ APPLICANT: Tabaska, Jack E.
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 16264
/ LENGTH: 624
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264
```

```
Query Match          5.6%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 6.7e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```
Qy 791 TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
Db 374 TTTTGGGCTAGTTGGAATCTCATTTTTCAGAGATTTTATTTCTTAAGAA 433
Qy 851 ATTAGTTATTTCTCTTATTAATAATGAAAAACATTGAAATTAAGTTGCCAGACTA 910
    |||||
Db 434 ATTAGTTATTTCTCTTGAAGAAATGAAATCCCTTGGAATTAAGTTCTAAACGA 493
Qy 911 GCCCTAAGT 920
    |||||
Db 494 GCCCTAATT 503
```

```
RESULT 8
US-10-425-115-141826
/ Sequence 141826, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 141826
/ LENGTH: 1326
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826
```

```
Query Match          5.6%; Score 73; DB 20; Length 1326;
```


Best Local Similarity 69.1%; Pred. No. 1.6e-05;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

QY 756 TGTCTAGATTCCTGTCGTAACAAATCTTGTGATTTTAAAGACGTAGTTGGCAACCT 815

Db 1090 TGAAGTTGGTGTGTCTGCTGATTTTCACTTTAACTGTAAGACTAATTTGGGAACCA 1149

QY 816 GTTCTTTCAAGAAATTTTGTATTTTCAAAAAAATTTAGTTATTTTCTTTATATAA 875

Db 1150 ATTT-TCACACTGATTTTCAATTTTCTTAAGAAAAATAGTTCATTTCCCTTGAGAAA 1208

QY 876 TAGAAAACTAGTAAATAATAGTTGCGAGACTAGCCCTAGAA 920

Db 1209 TAGAATCCCTTAGAAAAAATAAGTTTCCAAACTAGCCCTTAAAT 1253

RESULT 9
US-10-425-114-32493

Sequence 32493, Application US/10425114

Publication NO. US20040034888A1

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E.

APPLICANT: Tabaska, Jack E.

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(5313)B

CURRENT APPLICATION NUMBER: US/10/425, 114

NUMBER OF SEQ ID NOS: 2003-04-28

LENGTH: 2445

SEQ ID NO 32493

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI

US-10-425-114-32493

Query Match 5.6%; Score 72.8; DB 18; Length 2445;

Best Local Similarity 73.6%; Pred. No. 2.3e-05;

Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCCCTGTTCTT-TCAAAGATTTGATTTTCAAAAAA 850

Db 1623 TCTTAGGCTAGTTTGAAACCTTTTTCCTCAAAAGATTTTCAAAAGAAA 1682

QY 851 ATTAGTTATTTCTCTTTATTAATAATAGAAAAACATTAGAAAAATAGATTGCCAGCTA 910

Db 1683 ATTAGTTATTTCTCTTTATTAATAATAGAAAAATAGAAATCCCTTAAATAATAGTTTCAAACTA 1742

QY 911 GCCCTAGATGTTTCCCATTA 934

Db 1743 GCCCTTAATGTTTTCATGAA 1766

RESULT 10
US-10-425-115-83293/c

Sequence 83293, Application US/10425115

Publication NO. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425, 115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 83293

LENGTH: 2729

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(2729)

OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: MRT4577_175978C.1

US-10-425-115-83293

Query Match 5.6%; Score 72.8; DB 20; Length 2729;

Best Local Similarity 60.7%; Pred. No. 2.5e-05;

Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;

QY 792 TTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGATTTGATTTTCAAAAAA 851

Db 2523 TTAAAGGCTAGTTAGAA-CCTCATTTTCTTAAGATTTCTATTTTCCAGAAAA 2465

QY 852 TTAGTTATTTCTCTTTAATAATAAGAAAAACATTAGAAAAATAGATTGCCAGCTAG 911

Db 2464 ATATTTCATTTCCCTTGAATAATAGAAATACCTTGAAAAATAGATTCCAACTAG 2405

QY 912 CCTAGAAATGTTTCCCAATTAATCAATCACTGTATTAATTTTGGCCAGCCCAT 971

Db 2404 CCTTCAAAATTTCTCTCAAGTATTTCTTATAGAAAGTAAATGTAACATAGGAGCT 2345

QY 972 AAATTTAAACCGAACTGAAATGAGCGAAACCAATCTGA 1015

Db 2344 TGAGTATCATGAGCTTAAATGTAAAGGTAAATCAATGTTA 2301

RESULT 11
US-10-425-115-51470/c

Sequence 51470, Application US/10425115

Publication NO. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425, 115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 51470

LENGTH: 1203

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(1203)

OTHER INFORMATION: unsure at all n locations

OTHER INFORMATION: Clone ID: MRT4577_146936C.1

US-10-425-115-51470

Query Match 5.5%; Score 72.4; DB 20; Length 1203;

Best Local Similarity 76.1%; Pred. No. 2e-05;

Matches 102; Conservative 0; Mismatches 31; Indels 1; Gaps 1;

QY 787 ATTTTAAAGCTAGTTGGCAACCCCTGTTCTTCAAAAGATTTGATTTTCAAA 846

Db 1189 ACTTATTAAGGCTAGTGGGGA-CATATTTTCCAAAAAGATTTCTATTTCTTA 1131

QY 847 AAAATTAAGTTATTTCTTTATTAATAATAGAAAAACATTAGAAAAATAGAGTCCAG 906

Db 1130 GAAATTAAGTTATTTCTTTGAAAAATTAATAATCCGTGAAAAATAGAGTTCCAA 1071

QY 907 ACTAGCCCTAGAA 920

; ORGANISM: zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_147620C.1
US-10-425-115-52216

Query Match 5.5%; Score 71.8; DB 20; Length 2537;
Best Local Similarity 73.4%; Pred. No. 3.9e-05;
Matches 105; Conservative 0; Mismatches 37; Indels 1; Gaps 1;

QY	792	TTTAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAA	851
DB	1799	TCTTAGGGCTAGTTGGAACCT-TTTTCCCAAAAGATTTTCATTTTCAAAGAAAA	1857
QY	852	TTAGTTATTTCTCTTTAATAAATAGAAAAACATTAGAAAAATAGAGTGCAGACTAG	911
DB	1858	TTAGTTCATTTTCTTGAGAAAAATAGAAATCCCTTAAAAAATAGTGTTCAAACTAG	1917
QY	912	CCCTAGATGTTTCCCAATPAA	934
DB	1918	CCCTTAATGTTTTCATGAA	1940

Search completed: September 15, 2005, 20:45:54
Job time : 854.804 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 16:07:32 ; Search time 4088.83 Seconds
(Without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-2_COPY_1_1311

Perfect score: 1311
Sequence: 1 ccacgtgctcctcctacgaataa.....cgacacatgcttcctgaac 1311

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_nuc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	908	69.3	963	9	CC656933 OGMDO20TM
2	679	51.8	915	9	CG224225 OGIAG08TV
3	419.8	32.0	687	9	CC656939 OGMDO20TV
4	96	7.3	715	9	CG252571 CGAB057C
5	92	7.0	967	9	CL235046 ZMMBB057
6	88	6.7	814	9	CG048704 PU1L019TB
7	87.8	6.7	754	9	CG414922 ZMMBB025
8	86.8	6.6	950	8	CC439901 PUHVI15TB
9	86.6	6.6	652	8	CC384247 PUHOC67TB
10	86.6	6.6	797	8	CC400575 PUHLU61TD
11	86.6	6.6	820	8	CC400574 PUHLU61TB
12	85.4	6.5	471	9	CG103452 PUHBI19TB
13	85.2	6.5	765	9	CG082135 PUFOX12TD
14	85.2	6.5	781	9	CC630219 OGUCC53TV
15	85.2	6.5	815	9	CG349565 OGOFI183TH
16	85.2	6.5	834	9	CC630210 OGUCC53TH
17	85	6.5	789	8	CC433618 PUHHP17TD
18	84.6	6.5	1092	8	AL175666 Tetraodon
19	84	6.4	793	8	BZ816381 PUFBAB6TD
20	84	6.4	1078	9	CL957678 ZMMBH001
21	83.8	6.4	530	9	CG201774 PUICH24TB
22	83.6	6.4	722	9	CG333914 OGOAD14TH
23	83.6	6.4	722	9	CG333929 OGOAD14TV
24	83.6	6.4	781	9	CG034985 PUIGR68TB

25	83.6	6.4	861	8	BZ797976
26	83.6	6.4	865	8	CC430754 PUHEB05TB
27	81.8	6.2	861	9	CG102052 PUFTW94TD
28	81.8	6.2	947	8	CC435780 PUHNS02TD
29	81.8	6.2	981	8	BZ784278 PUHFV09TD
30	81.8	6.2	1017	9	CC620594 OGUCC26TV
31	81.6	6.2	649	8	CC613918 OGIAG88TV
32	81.6	6.2	733	8	BZ778636 1h02110.9
33	81.6	6.2	925	9	CG071791 PUHUK02TB
34	81.6	6.2	960	8	BZ676889 PUHIG17TD
35	81.6	6.2	994	8	CC003943 PUJUN65TD
36	81.6	6.2	1016	9	CL996481 ZMMBH000
37	81.6	6.2	1022	9	CL984151 ZMMBH000
38	81.4	6.2	999	8	CC385762 PUHFK14TD
39	81.2	6.2	756	9	CG221693 OGMGO04TH
40	81	6.2	1101	9	AL078714 Drosophila
41	80.6	6.1	699	8	BZ996930 PUHGR50TB
42	80.6	6.1	832	9	CC676387 OGMAS81TH
43	80.6	6.1	839	9	CG246849 OGMAL22TH
44	80.6	6.1	843	9	CG254365 OGMFV41TH
45	80.6	6.1	852	8	CC385881 PUHMD82TD

ALIGNMENTS

RESULT 1
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGMDO20TM ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.

ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.

SOURCE
ORGANISM Zea mays

REFERENCE
AUTHORS Whitehead, C.A., Quackenbush, J., Van Aken, S., Uteerback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
1 (bases 1 to 963)
clade; Panicoideae; Andropogoneae; Zea.

TITLE
JOURNAL
COMMENT
Other GSSs: OGMDO20TV
Unpublished (2002)
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0554D15"
/clone_1ib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-7 Site 1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 69.3%; Score 908; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 7.8e+180; Indels 0; Gaps 0;
Matches 908; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
404 ACCAGCCATGTCCTAGTCCACTATTCGACCTACCCAGACAAAGATTTAAAAAATTAACC 463
1 ACCAGCCATGTCCTAGTCCACTATTCGACCTACCCAGACAAAGATTTAAAAAATTAACC 60

QY 1221 TGCGTATTCGAAACGAAGATACCTCCAAACAAATCCATCTTACTCATGCAACT 1280
 |||||
 DB 600 TGCGTATTCGAAACGAAGATACCTCCAAACAAATCCATCTTACTCATGCAACT 659

QY 1281 CCATGCAACACGACATATGTTCTCGAAC 1311
 |||||
 DB 660 CCATGCAACACGACATATGTTCTCGAAC 690

RESULT 3
 CC656939 687 bp DNA linear GSS 19-JUN-2003
 LOCUS OGMDQ20TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0554D15,
 DEFINITION genomic survey sequence.
 ACCESSION CC656939
 VERSION CC656939.1 GI:32060231
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 687)

REFERENCE
 Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGMDQ20TM
 Contact: Cathy Whitefaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitefaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..687
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBma0554D15"
 /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 32.0%; Score 419.8; DB 9; Length 687;
 Best Local Similarity 99.5%; Pred. No. 1.6e-77;
 Matches 421; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 889 GAAAAATAGAGTCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGG 948
 |||||
 DB 687 GAAAAATAGAGTCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCAATCACTGG 628

QY 949 TATAATTAATTTGGCCAGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1008
 |||||
 DB 627 TATAATTAATTTGGCCAGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 568

QY 1009 AATCTGAGCTATTTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
 |||||
 DB 567 AATCTGAGCTATTTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508

QY 1069 AGTCATTTGCCGAGATGAGGTTTGGCAAGATAGCCAGTATCATAGCTATG 1128
 |||||
 DB 507 AGTCATTTGCCGAGATGAGGTTTGGCAAGATAGCCAGTATCATAGCTATG 448

QY 1129 GTGCTTAAGTCAAGTTCCGAGCTCTGTCATCTTCAATGAGCATTAATGCTTTGTT 1188
 |||||
 DB 447 GTGCTTAAGTCAAGTTCCGAGCTCTGTCATCTTCAATGAGCATTAATGCTTTGTT 388

QY 1189 CAACGGTTCGTTGTCATGTCGAAGCCTTGCTATTCGAAACGAAGATACCTTA 1248
 |||||
 DB 387 CAACGGTTCGTTGTCATGTCGAAGCCTTGCTATTCGAAACGAAGATACCTTA 328

QY 1249 CTCGCAACAAATCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCTG 1308
 |||||
 DB 327 CTCGCAACAAATCATCTTACTCATGCAACTTCATGCAACACGACATATGTTTCTG 268

QY 1309 AAC 1311
 |||||
 DB 267 AAC 265

RESULT 4
 CG252571 715 bp DNA linear GSS 25-AUG-2003
 LOCUS OG4B05TC_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMBma0809B10,
 DEFINITION genomic survey sequence.
 ACCESSION CG252571
 VERSION CG252571.1 GI:34154661
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 715)

REFERENCE
 Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Contact: Cathy Whitefaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitefaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 1..715
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBma0809B10"
 /note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 7.3%; Score 96; DB 9; Length 715;
 Best Local Similarity 80.1%; Pred. No. 1.2e-09;
 Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;

QY 135 TCTCCACCTCGCGTGCATCTTCTTGATGCGGTGTTCCATCTGACGAGGCCCA 194
 |||||
 DB 356 TCACACACCGCGTCTTGATGCGGTGTTCCATCTGACGAGGCCCA 297

QY 195 TCAGACACCTTTCCGAGACCCATCAAGGCTTTTCGATGGCCACAGAGACTATCGGG 254
 |||||
 DB 296 TAAAGCACTGTCGAGAACCCATTAAGGCTTTTCGATGGCCACAGAGACTATCGGG 238

QY 255 TCGGTGATCCAGGGATATATGTCCCCCAATC 290
 |||||
 DB 237 TCGGTGATCCAGGGATATATGTCCCCCAATC 202

RESULT 5
 CL235046 967 bp DNA linear GSS 15-JAN-2004
 LOCUS ZMBBb0575001r_ZMMBB (HindIII) Zea mays genomic clone
 DEFINITION

ACCESSION	ZMMBBD0575001.3', genomic survey sequence.		
VERSION	CL235046		
KEYWORDS	CL235046.1 GI:40891729		
SOURCE	GSS.		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 967) Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C., Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.		
TITLE	Sequencing of the maize genome at PGIR (2003c)		
JOURNAL	Unpublished (2003)		
COMMENT	Contact: Bharti,A.K. Dr. Joachim Messing's lab The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University 190 Frelinghuysen Road, Piscataway, NJ 08854, USA Tel: 732 445 3801 Fax: 732 445 5735 Email: bharti@waksman.rutgers.edu Seq primer: SP6 Clases: BAC ends High quality sequence start: 64. Location/Qualifiers 1..967 /organism="Zea mays" /mol_type="genomic DNA" /culivar="B73" /db_xref="taxon:4577" /clone="ZMMBBD0575001" /lab_host="E. coli DH10B" /clone_1ib="ZMMBBD (HindIII)" /note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"		
ORIGIN			
Query Match	7.0%	Score 92;	DB 9; Length 967;
Best Local Similarity	78.6%	Pred: NO. 8.5e-09;	
Matches 110; Conservative	0;	Matches 30; Indels	0; Gaps 0;
Oy	150	TGCACATTTCTTGATGTGCGGTGTCCTCATCTGACGAGGCGCCATCAGACACCTTTGG	209
Db	827	TGCGCCTTCTTGATACCGCGGTTCCCGTCTGATTAGGCCATTCAGACACCTTTGG	886
Oy	210	GACACCCATCAAGGGCTTTGGATGAGCCACGAGACGATCGGCTGCGTGTGATCCAG	269
Db	887	GTGGCGCCATCAAGTGCTTTGGATGATGCCACGGAGTATCCGGCGCGTGACCTGG	946
Oy	270	GGATTATATGTCCTCCCAAT	289
Db	947	GGATTATGTCCTCCCAATAGT	966
RESULT 6			
CG048704			
LOCUS	CG048704	814 bp	DNA linear GSS 19-AUG-2003
DEFINITION	PULIO19TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0611C13,		
ACCESSION	CG048704		
VERSION	CG048704.1 GI:33920884		
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 814) Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Benwick,A., Frazer,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.		
TITLE	Maize Genomics Consortium		
JOURNAL	Unpublished (2003)		

COMMENT	Other GSSs: F01L01.9TD	Contact: Cathy Whitelaw
	9712 Medical Center Drive, Rockville, MD 20850, USA	
	Tel: 301-838-5843	
	Fax: 301-838-0208	
	Email: whitelaw@cigr.org	
	Seq primer: TR	
	Class: sheared ends.	
FEATURES	Location/Qualifiers	
SOURCE	1. 814	
	/organism="Zea mays"	
	/mol_type="genomic DNA"	
	/strain="973"	
	/db_xref="taxon:4577"	
	/clone="ZMMBR0611C13"	
	/clone_1lb="ZM_0.6-1.0_KB"	
	/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high	
	Cor selected genomic DNA library"	
ORIGIN		
Query Match	6.7%; Score 88; DB 9; Length 814;	
Best Local Similarity	62.7%; Pred. No. 5.8e-08;	
Matches 153; Conservative	0; Mismatches 90; Indels 1; Gaps 1;	
Qy	713 TTTATTAATGACTATTAAGTCATTTTATATATAAGCAGCATGTCGTAGATTCGTGTT 772	
Db	487 TTTTAGAGGACGTTGCTAGATGATGGAAGAATATATAGATAATATCTTTAGAGATGTTA 546	
Qy	773 CAAAAATCTTCGATTTTTTTTAAAGCTAGTTGGCAACCCGTGTTCTTCAAGAATT 832	
Db	547 TAAAGGACAGAAACAATTCCTTAAGAGCTAGTTGGTAATCCCATTTTTCCAAAGATT 606	
Qy	833 TTGATTTTTTCAAAA-AAAATGATTATTTCTCTTAAATTAATAATGAAAAACATTAGAA 891	
Db	607 TTCATTCTCACAAGGAGAAAATAGTTATTTCCCTTTGTAATAATGTATCACTCGGAA 666	
Qy	892 AAATAGATTGCCAGACTAGCCCTAGATGTTTCCCAATTAATTAACAATCACTGTGTA 951	
Db	667 AAATAGGTTCTCAAACTAGCCCTAAGGTTATATCCGATTAATCCCTATCTCATTTCT 726	
Qy	952 AATT 955	
Db	727 ATTT 730	
RESULT 7		
CG414922/c	754 bp DNA 1linear GSS 08-SEP-2003	
LOCUS	ZMMBB0290L09.r ZMMBB Zea mays genomic clone ZMMBB0290L09 3',	
DEFINITION	genomic survey sequence.	
ACCESSION	CG414922	
VERSION	CG414922.1 GI:34505144	
KEYWORDS	GSS.	
SOURCE		
ORGANISM	Zea mays	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	
REFERENCE	1 (bases 1 to 754)	
AUTHORS	Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bhatti, A.K., Messing, J. and Wing, R.	
TITLE	Sequencing of the maize genome	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Rod Wing	
	Arizona Genomics Institute	
	University of Arizona	
	Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ	
	85721-0088, USA	
	Tel: 520 626 3867	
	Fax: 520 621 9288	
	Email: http://genome.arizona.edu	
	PCR Primers	

Db 461 TATTTTTCAGTGAATTAAGTTATTTTTCATGCGAATAAGAAATTAATGGAAAA 520
QY 894 ATAGAGTGGCAGACTAGCCCTAGA 918
Db 521 ATTAGTTTCCAACTAGCCCTAAA 545

RESULT 10
CC400575/c
LOCUS CC400575 797 bp DNA linear GSS 19-MAY-2003
DEFINITION PUHL061TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBR1480L01,
genomic survey sequence.
ACCESSION CC400575
VERSION CC400575.1 GI:30880665
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 797)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHL061TD
Contact: Cathy WhiteLaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source location/Qualifiers
1..797
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBR1480L01"
/note="Vector: PCR4-TOPO; Site: 1; EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.6%; Score 86.6; DB 8; Length 797;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;
Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCGATGCTAGATTCGTT 772
Db 577 TTTAGGAGCGTTGTAAGATGAAAGATATAGAAAATATCTTTTAGAGATGTTA 518
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTTCAAAGATT 832
Db 517 TAAAGACAAGAAAATTTTATAGAGCTAGTTGGTAATCCCATTTTTCAAAGATT 458
QY 833 TTGATTTTTCAAA-AAAATTAGTTATTTCTTTATTAATGAAAACCTTAGAA 891
Db 457 TTCAATCTCACAGGAGAAAATAGTTATTTCCCTTTTGAATGTAATCATCGGAA 398
QY 892 AATAGAGTGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAACATCACTGTAT 951
Db 397 AATTAAGATTCTCAAACTAGCCCTAAGTTATATCCGATTAATCCCTATCTCTCT 338
QY 952 AATTATTTGGCCAGCCCATTAATTAATTTTAAC 984
Db 337 ATTCAAAATTCACCTTGCAAAATAGTGTATC 305

RESULT 11

CC400574
LOCUS CC400574 820 bp DNA linear GSS 20-AUG-2003
DEFINITION PUHL061TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBR1480L01,
genomic survey sequence.
ACCESSION CC400574
VERSION CC400574.1 GI:30880664
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 820)
WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUHL061TD
Contact: Cathy WhiteLaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteLaw@tigr.org
Seq primer: TF
Class: sheared ends.

FEATURES
source location/Qualifiers
1..820
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMBR1480L01"
/note="Vector: PCR4-TOPO; Site: 1; EcoRI; 0.6-1.0 kb high
cot selected genomic DNA library"

ORIGIN
Query Match 6.6%; Score 86.6; DB 8; Length 820;
Best Local Similarity 59.7%; Pred. No. 1.1e-07;
Matches 163; Conservative 0; Mismatches 109; Indels 1; Gaps 1;

QY 713 TTTATTAATGACTATTAAGTCAATTTTATATAGAGCAGCGATGCTAGATTCGTT 772
Db 374 TTTAGGAGCGTTGTAAGATGAAAGATATAGAAAATATCTTTTAGAGATGTTA 433
QY 773 CAAAAATCTTCTGATTTTATAGAGCTAGTTGGCAACCTGTTCTTTCAAAGATT 832
Db 434 TAAAGACAAGAAAATTTTATAGAGCTAGTTGGTAATCCCATTTTTCAAAGATT 493
QY 833 TTGATTTTTCAAA-AAAATTAGTTATTTCTTTATTAATGAAAACACTTAGAA 891
Db 494 TTCAATCTCACAGGAGAAAATAGTTATTTCCCTTTTGAATGTAATCATCGGAA 553
QY 892 AATAGAGTGGCAGACTAGCCCTAGATGTTTCCCAATTAATTAACATCACTGTAT 951
Db 554 AATTAAGATTCTCAAACTAGCCCTAAGTTATATCCGATTAATCCCTATCTCTCT 613
QY 952 AATTATTTGGCCAGCCCATTAATTAATTTTAAC 984
Db 614 ATTCAAAATTCACCTTGCAAAATAGTGTATC 646

RESULT 12
CG103452
LOCUS CG103452 471 bp DNA linear GSS 20-AUG-2003
DEFINITION PUB0619TB_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMBR1480L01,
genomic survey sequence.
ACCESSION CG103452
VERSION CG103452.1 GI:33985746
KEYWORDS GSS.
SOURCE Zea mays

ORGANISM	Zeae mayas
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 471)
AUTHORS	Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Renick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
TITLE	Maize Genomes Consortium
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: PUFB19TD Contact: Cathy Whitelaw
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: TR Class: sheared ends.
FEATURES	Location/Qualifiers
SOURCE	1..471 /organism="Zea mays" /mol_type="genomic DNA" /strain="873" /db_xref="taxon:4577" /clone="ZMBMTA0628D13" /clone_1lb="ZM.0.6.1.0 KB" /note="Vector: PCR4-TOPO, Site 1: EcoRI; 0.6-1.0 kb high COT selected genomic DNA library"
ORIGIN	
Query Match	6.5%; Score 85.4; DB 9; Length 471;
Best Local Similarity	69.5%; Pred. No. 2e-07;
Matches 116; Conservative	0; Mismatches 51; Indels 0; Gaps 0;
Qy	789 TTTTAAAGCTACTTTGGCAACCTGTTCTTCAAGAATTGTGATTTTGCAGAAA 848
Db	216 TTTCTTAGAGCTACTTTGGAAATCCATTTTTCAGGATTTACATTTTTCAGGG 275
Qy	849 AAATTAGTTATTTTCTTTTATAAATAGAAACACTTAGAAAATAGATTGCCAGAC 908
Db	276 AAATTAGTTATTTTCTTTGAAAATAGAAATCAGATCGGGAATATGTGTCCCAAC 335
Qy	909 TAGCCCTAGAAATGTTTCCCATTAATTAACATCATCTGTATTAAT 955
Db	336 TAGCCCTAGAAATGTTTCTTAAAAAGATGATCAATGATTAATTAAT 382
RESULT 13	
CG082135/c	765 bp DNA linear GSS 20-AUG-2003
LOCUS	PUFOX12TD ZM.0.6.1.0 KB Zea mays genomic clone ZMBMTA0712B24,
DEFINITION	genomic survey sequence.
ACCESSION	CG082135
VERSION	CG082135.1 GI:33964429
KEYWORDS	GSS.
SOURCE	
ORGANISM	Zea mays
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE	1 (bases 1 to 765)
AUTHORS	Whiteclaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Renick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
TITLE	Maize Genomes Consortium
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: PUFOX12TB Contact: Cathy Whitelaw
TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208

Email: whitejaw@icigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 1..765
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMB7A0712B24"
 /clone_1id="ZM 0.7-1.5 KB"
 /note="Vector: pCR4-TOP0; Site 1: EcoRI; 0.6-1.0 kb high
 cor selected genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 765;
 Best Local Similarity 71.8%; Pred. No. 2.2e-07;
 Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

QY 787 ATTTTATAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTTTCAAA 846
 DB 255 ATCTTTCAATAGACTAGTTGGGAACCTTAATTT-TCACGGGAGATTTTATTTTTCAT 197
 QY 847 AAAAATTAGTTATTTTCTCTTATTAATAATGAAACCTTAGAAATATAGATTGCGAC 906
 DB 196 GAAAATTAAGTTATTTTCTCTTGGAAATAATAATCATCTTGAGAAATATAGTTTCCAA 137
 QY 907 ACTAGCCCTAGATATGTTTCCCAATTAATTAATCATCATCTGATTAATTAATTTT 960
 DB 136 ACTAGCCCTAATAGTATTTTATTAATAATGAGGAAAAATTCGTCATTTTTCATTG 83

RESULT 14
 CC630219/c
 LOCUS
 DEFINITION OGUUGS3TV_ZM_0.7_1.5_KB_Zea_mays_genomic_clone_ZMMB7A0406J09,
 genomic survey sequence.
 CC630219
 CC630219.1 GI:32003354
 GSS.
 SOURCE
 ORGANISM
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 781)
 WhiteJaw C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Renick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Citek, R.W., Nunez, A., Robbins, D. and Lakey, N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGUUGS3TH
 Contact: Cathy WhiteJaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitejaw@icigr.org
 Seq primer: TP
 Class: sheared ends.
 Location/Qualifiers
 1..781
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMB7A0406J09"
 /clone_1id="ZM 0.7-1.5 KB"
 /note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 781;

Search completed: September 16, 2005, 08:08:42
Job time : 4089.83 secs

Best Local Similarity 71.8%; Pred. No. 2.2e-07;
Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGACGAGTTGGCAACCGTCTTTCACAAAGATTGATTTTCAAA 846
    |||||
DB 665 ATCTTTCTAATGACTAGTTGGGACCTTATTT-TCTCAGGGATTTTATTTTCAAT 607
    |||||
QY 847 AAAAATTAGTTTCTTCTTATATAAAGAAAACCTTAGAAAAATAGAGTTGCCAG 906
    |||||
DB 606 GAAAATTAGTTTATTTCTCTTGAGAAAATATTAATCATTGAGAAAATATAGTTCCAA 547
    |||||
QY 907 ACTAGCCCTGAGATGTTTCCCATATAATACATCATGCTATATATTATTG 960
    |||||
DB 546 ACTAGCCCTGAGATTTTATATAATAGAGAAAATGCTATTTTTCATTG 493
    |||||
```

RESULT 15

CG349565 815 bp DNA linear GSS 26-AUG-2003
LOCUS CG349565/c
DEFINITION OG0F183TH_ZM_0_7_1_5_Zea_mays_genomic_clone_ZMMBMA0705N21,
genomic survey sequence.

ACCESSION CG349565
VERSION CG349565.1 GI:34266831
KEYWORDS GSS.

SOURCE

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 815)

REFERENCE

AUTHORS Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
White, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Whitelaw, C.A., Nunberg, A., Robb, D. and Lakey, N.
Citek, R.W., Nunberg, A., Robb, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG0F183TV
Contact: Cathy Whitelaw

TITLE

JOURNAL
COMMENT
TIQR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Classes: sheared ends.

FEATURES

source
location/Qualifiers
1..815
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 6.5%; Score 85.2; DB 9; Length 815;
Best Local Similarity 71.8%; Pred. No. 2.2e-07;

Matches 125; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

```
QY 787 ATTTTAAAGACGAGTTGGCAACCGTCTTTCACAAAGATTGATTTTCAAA 846
    |||||
DB 173 ATCTTTCTAATGACTAGTTGGGACCTTATTT-TCTCAGGGATTTTATTTTCAAT 115
    |||||
QY 847 AAAAATTAGTTTCTTCTTATATAAAGAAAACCTTAGAAAAATAGAGTTGCCAG 906
    |||||
DB 114 GAAAATTAGTTTATTTCTCTTGAGAAAATATTAATCATTGAGAAAATATAGTTCCAA 55
    |||||
QY 907 ACTAGCCCTGAGATGTTTCCCATATAATACATCATGCTATATATTATTG 960
    |||||
DB 54 ACTAGCCCTGAGATTTTATATAATAGAGAAAATGCTATTTTTCATTG 1
    |||||
```


Db	181	CTGACCGGAGGCCCATCAGACACCTTTGGGACACCCATCAAGGGCTTTGGATGGGCCA	240
Qy	241	CGAGACGTATCGGGTCTGGTGTGATCAGAGGGATATATGTCTCCCAACAAATGTCACCTATA	300
Db	241	CGAGACGTATCGGGTCTGGTGTGATCAGAGGGATATATGTCTCCCAACAAATGTCACCTATA	300
Qy	301	TTATTAATCTTTAGATATATTTATTTTGGAAAAAATACAAACTTATACCTTTTGGTA	360
Db	301	TTATTAATCTTTAGATATATTTATTTTGGAAAAAATACAAACTTATACCTTTTGGTA	360
Qy	361	GGGCGCTCAGACATAGATTTTTGCTTAAGGGCCAGAAATCGAGACACAGCCATGTCATGTG	420
Db	361	GGGCGCTCAGACATAGATTTTTGCTTAAGGGCCAGAAATCGAGACACAGCCATGTCATGTG	420
Qy	421	TCACATATTTGGACATACCCAGAACAAATTTAAAAAATMACAAAGTACTAATCCACT	480
Db	421	TCACATATTTGGACATACCCAGAACAAATTTAAAAAATMACAAAGTACTAATCCACT	480
Qy	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACAGATCCCTTAAAAAA	540
Db	481	CGAAAGCTATCATGTATGTTTAAAGAAACATCTATTAAACACAGATCCCTTAAAAAA	540
Qy	541	CAAGCATATTTGAAAAGACAAATTAATGTTTACAGTTTACAAACATCTAAGAGCGACAA	600
Db	541	CAAGCATATTTGAAAAGACAAATTAATGTTTACAGTTTACAAACATCTAAGAGCGACAA	600
Qy	601	TTATTCGAAAGGTAGCTATGAGGTGAGATTTTCTTTTCAATCTGTGTAATTTGT	660
Db	601	TTATTCGAAAGGTAGCTATGAGGTGAGATTTTCTTTTCAATCTGTGTAATTTGT	660
Qy	661	ATTGTTTTATATACATTTTCTTCTCTTACAAATAGATTTTCTTCGATTTTATATAA	720
Db	661	ATTGTTTTATATACATTTTCTTCTCTTACAAATAGATTTTCTTCGATTTTATATAA	720
Qy	721	ATGACTATPAAGTCAATTTTATATATAGACACGATGTCGTAGATCTTCGTCAAAAAATC	780
Db	721	ATGACTATPAAGTCAATTTTATATATAGACACGATGTCGTAGATCTTCGTCAAAAAATC	780
Qy	781	TTTTCGATTTTTTAAAGACTAGTTTGCCAAACCCGTTCCTTTCGAAAGAAATTTGATTTT	840
Db	781	TTTTCGATTTTTTAAAGACTAGTTTGCCAAACCCGTTCCTTTCGAAAGAAATTTGATTTT	840
Qy	841	TTCAAAAAAATTAAGTTATTTTCTCTTTATATAATAGAAAAACATTAGAAAAATAGACT	900
Db	841	TTCAAAAAAATTAAGTTATTTTCTCTTTATATAATAGAAAAACATTAGAAAAATAGACT	900
Qy	901	TGCCAGACTAGCCCTCAGAAATGTTTTCCCAATPAATTAACATCAGTGTATATATATG	960
Db	901	TGCCAGACTAGCCCTCAGAAATGTTTTCCCAATPAATTAACATCAGTGTATATATATG	960
Qy	961	GCCAGCCCCCATPAATTAATTTTAAACCGAAACTGAAATCGAGCGAAACCAATCTGAGCTAT	1020
Db	961	GCCAGCCCCCATPAATTAATTTTAAACCGAAACTGAAATCGAGCGAAACCAATCTGAGCTAT	1020
Qy	1021	TTCTCTAATTAAGTAAAAAGGAGAGAGAGGAAATACAGTTTATAGTCATTTGTC	1080
Db	1021	TTCTCTAATTAAGTAAAAAGGAGAGAGAGGAAATACAGTTTATAGTCATTTGTC	1080
Qy	1081	TGAGATGTGCGGTTTGGACAGATAGCCACCGTATCATACCTCATAGTGTCCCTACGCTA	1140
Db	1081	TGAGATGTGCGGTTTGGACAGATAGCCACCGTATCATACCTCATAGTGTCCCTACGCTA	1140
Qy	1141	GGTTTCGAGGCTCTCGTGTCAATCTCAATGCGATCTCAATGCTTTGTCACACGTTGCTC	1200
Db	1141	GGTTTCGAGGCTCTCGTGTCAATCTCAATGCGATCTCAATGCTTTGTCACACGTTGCTC	1200
Qy	1201	TTGTTCCATGCTCAGAGCTTGGCTTAATCTGAAACCAAGAGATACCTACCTCCCAACAT	1260
Db	1201	TTGTTCCATGCTCAGAGCTTGGCTTAATCTGAAACCAAGAGATACCTACCTCCCAACAT	1260
Qy	1261	CCATTTTACTCATGCACTTCCATGCAACACGACATATGTTTTCTGAAACCAATTCATTT	1320
Db	1261	CCATTTTACTCATGCACTTCCATGCAACACGACATATGTTTTCTGAAACCAATTCATTT	1320

QY	1321	AAAGTACAAACAGCTAGCGTTCTCCCGGCTAAGCTTCCTCTCTCTCTGCCGATCTTTT	1380
Db	1321	AAAGATCAACAACAGCTAGCGTTCTCCCGGCTAAGCTTCCTCTCTCTGCCGATCTTTT	1380
QY	1381	CGTCACGACCATG 1394	
Db	1381	CGTCACGACCATG 1394	
RESULT 2			
BD062176			
LOCUS	BD062176	1394 bp	DNA
DEFINITION	Male tissue-preferred regulatory region and method of using same.		
ACCESSION	BD062176		
VERSION	BD062176.1	GI:22607781	
KEYWORDS	JP 2001520523-A/1.		
SOURCE	JP 2001520523-A/1.		
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and		
AUTHORS	Kendall,T.L.		
TITLE	Male tissue-preferred regulatory region and method of using same		
JOURNAL	Patent: JP 2001520523-A 1 30-OCT-2001;		
COMMENT	PIONEER HI BRED INTERNATIONAL INC		
	PN JP 2001520523-A/1		
	PD 30-OCT-2001		
	PF 19-JUN-1998 JP 1999504910		
	PR 23-JUN-1997 US 08/880499		
	PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,		
	PI TIMMY L KENDALL		
	PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC		
	,C07K14/34,C12Q1/68,		
	PC A01H5/00		
	CC Strandedness: Single;		
	CC Topology: Linear;		
FEATURES	key	Location/Qualifiers.	
source	1..1394	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
	Query Match	100.0%; Score 1394; DB 6; Length 1394;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 1394; Conservative	0; Mismatches	0; Indels
QY	1	CCATGCGTCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTTCTTAGGGTCC	60
Db	1	CCATGCGTCTCTATGAAAAAGATGATACATGTGCTATATCCGTTTTCTTAGGGTCC	60
QY	61	CTTCTTCGCTTATCTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTCATGAT	120
Db	61	CTTCTTCGCTTATCTACTGACTGAATCGGGGTTACAAAAAACTTCCACGGGTCATGAT	120
QY	121	CTTCATGTTCCACTTCTCCACCTCGGGGTTGACATTTCTTGATGTGCGTGTCTCCAT	180
Db	121	CTTCATGTTCCACTTCTCCACCTCGGGGTTGACATTTCTTGATGTGCGTGTCTCCAT	180
QY	181	CTGACCGAGGCCCATGACACCTTTTGGGACACCCATCAAGGGCTTTTGGATGGCCCA	240
Db	181	CTGACCGAGGCCCATGACACCTTTTGGGACACCCATCAAGGGCTTTTGGATGGCCCA	240
QY	241	CGAAGCGTATCGGTCGTGTGATCCAGAGGGATATATGTCCCCACATCGTCACTATA	300
Db	241	CGAAGCGTATCGGTCGTGTGATCCAGAGGGATATATGTCCCCACATCGTCACTATA	300
QY	301	TTATTTATCTTTAGATATATTTATTTTGGAAAAATTAACAACTTATACTTTTGTGTA	360
Db	301	TTATTTATCTTTAGATATATTTATTTTGGAAAAATTAACAACTTATACTTTTGTGTA	360

FEATURES	Location/Qualifiers
source	Location/Qualifiers
1..1394	
/organism="Homo sapiens"	
/mol_type="genomic DNA"	
/db_xref="taxon:9606"	

Query Match	Score	DB	Length
100.0%	1394	6	1394

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CCATGGTCTCTATGAAGAAGTAGACCAATGTCTAATATCCGTTTCTTAGGTC 60

Db 1 CCATGCTGCTCTATGAAAAAGTAGTACAAATGTCCTATATCCGTTTCTTAGGCTCC 60

61 CTTCTTCGCGCTATTACTGACTGAAATCGGGGTTACAAAAAATTCACCGGGTCATGAT 120

Db 61 CTTCTTCTGCCATTACTGACTGATGGGGTTACAAAACTTCCACGGGTGCATGAT 120

121 CTCGATGTTCACTTCTCCACCTGGCGTTGCATTTCTTGGATGTCGGTGTCCCAT 180

Db 121 CTCGATGTTCCACTTCTCCACCTGGCGTTGCACATTTCTTGGATGTCGGTGTGCCAT 180

QY 181 CTGACCGAGGCCCATCAGACACTTTCGGGACACCCATCAAGGGCTTTCGGATGGCCCA 240

Db 181 CTGACCGAGGCCCATCAGACACCTTTCCGGACACCCATCAAGGCGCTTTCCGATGCGCCA 240

241 CGAGAGTATCGGGTCTGTGTAATCCAGGGGATATATGTCCCCCAATCGTCACCTATA 300

Db 241 CGAGACGTATCGGGTCGTGTGATCCAGGGGATATGTGCCCCCAATCGTCACCTATA 300

301 TTATTATCTTAGAATTAATTAATTTTGGAAAAATACCAACATATACCTTGTGTA 360

DB 301 TTAATATCTTAGAATATATTTAAATTTTGGAAAAATACCAACTATATACITTTGGTGA 360

QY 361 GGGGCTGAGATGATTTGGCTTAGGGCCCGAATGCGAGACGACGATGCTAGTG 420
 DB 361 GGGGCTGAGATGATTTGGCTTAGGGCCCGAATGCGAGACGACGATGCTAGTG 420
 QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATAACCAAGTAATCACT 480
 DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTAAAAAATAACCAAGTAATCACT 480
 QY 481 CGAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
 DB 481 CGAAGCTATCATGTATGTTTAAAGAACATCTATTAAAAACGATCTCTTAAAAAA 540
 QY 541 CAAGCAATTTCCAGAAAGACAAATTTGTTACGTTTCAAAACATCTAAGACGACAA 600
 DB 541 CAAGCAATTTCCAGAAAGACAAATTTGTTACGTTTCAAAACATCTAAGACGACAA 600
 QY 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTGTT 660
 DB 601 TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTGTT 660
 QY 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGAGTATTTTCCGATTTTATAAA 720
 DB 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGAGTATTTTCCGATTTTATAAA 720
 QY 721 ATGACTATAAGTCAATTTTATATAGACGACGATGCTGATGCTGCTCAAAATC 780
 DB 721 ATGACTATAAGTCAATTTTATATAGACGACGATGCTGATGCTGCTCAAAATC 780
 QY 781 TTTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
 DB 781 TTTCTGATTTTTTAAAGAGTAGTTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
 QY 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATAGAAACACTTAGAAAAATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTTCTTTTAAATAGAAACACTTAGAAAAATAGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATTAATTCATCACTGCTATTAATTG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCATTAATTCATCACTGCTATTAATTG 960
 QY 961 GCCAGCCCCATTAATTTTAAACCGAATCGAATGACGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCCATTAATTTTAAACCGAATCGAATGACGAAACCAATCTGAGCTAT 1020
 QY 1021 TTTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAACTGATGCTCC 1080
 DB 1021 TTTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAACTGATGCTCC 1080
 QY 1081 TGAAGATGTGGGTTTGGCAAGATGACACCGTATATGCTCATAGGCTAGCTCA 1140
 DB 1081 TGAAGATGTGGGTTTGGCAAGATGACACCGTATATGCTCATAGGCTAGCTCA 1140
 QY 1141 GGTTCGGACAGCTCTCGTGTATCTCAATGCAATGCAATGCTGTTCAACCGTCTGTC 1200
 DB 1141 GGTTCGGACAGCTCTCGTGTATCTCAATGCAATGCAATGCTGTTCAACCGTCTGTC 1200
 QY 1201 TTGTTCCATGTCGCAAGCCTTGCCTATTCTGAACCAAGAGATACCTATCCCAACAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCTTGCCTATTCTGAACCAAGAGATACCTATCCCAACAT 1260
 QY 1261 CCATCTTACTACGCAACTTCCATGCAAAACGACATATGTTTCTGGAACCAATGCTAT 1320
 DB 1261 CCATCTTACTACGCAACTTCCATGCAAAACGACATATGTTTCTGGAACCAATGCTAT 1320
 QY 1321 AAAGATCAACAAGCTAGCTGTTCTCCGCTAGCTGCTCTCTCTGCGAATCTTTTT 1380
 DB 1321 AAAGATCAACAAGCTAGCTGTTCTCCGCTAGCTGCTCTCTCTGCGAATCTTTTT 1380
 QY 1381 CGTCCACCAACATG 1394
 DB 1381 CGTCCACCAACATG 1394
 QY 1394 CGTCCACCAACATG 1394
 DB 1394 CGTCCACCAACATG 1394

RESULT 3
 AF360356 3343 bp DNA linear PLN 12-MAY-2001
 LOCUS Zea mays male fertility protein (Ms45) gene, complete cds.
 DEFINITION
 ACCESSION AF360356
 VERSION AF360356.1 GI:114028756
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCB
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE
 1 (bases 1 to 3343)
 Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
 Cloning of Ms45, a gene required for male fertility from Zea mays
 Unpublished
 2 (bases 1 to 3343)
 Fox,T.W., Trimmell,M.R. and Albertsen,M.C.
 Direct Submission
 Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
 Hi-Bred Intl, Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
 IA 50131-1004, USA
 FEATURES
 source
 1..3343
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /cultivar="B73"
 /db_xref="taxon:4577"
 /chromosome="9L"
 <1392..>2942
 /gene="Ms45"
 join(<1392..1768,1898..2182,2280..2447,2534..>2942)
 /product="male fertility protein"
 join(1392..1768,1898..2182,2280..2447,2534..2942)
 /gene="Ms45"
 /codon_start=1
 /product="male fertility protein"
 /protein_id="AAK52489.1"
 /db_xref="GI:14028757"
 /translation="MEKRLQWRGRDGIVOYPHLFPALALALVADPFGSLAEV
 DYRVRKELAPYGEVMSWPRDASRGRLEPYGEVSPESIRPDLOGRPGVGLA
 DRRVVRMGEERGMETFAVMPDMSBEBCANCVNSTRKQHEKEFCGRPLGRHGE
 TGEIVADAVYGLMVVGSGVASSVAEADSDPIRFANDLVHNRGSVFPTDSNRK
 SRKDHNTLLEBEGRLRYDPEYSGVYVYKGLVFPNGVVISDHPFLRSETTNC
 RIMRYMLEBPRAVEFANLPSPDNVNSGRQFWVALDCRTPADQVFAKRPMLR
 TLVFPRLSLKYLTKAARRMHTVALLDGSRVVEVLEDRHEWKLVSREVRGRK
 LMIGVANHNTIIPYLED"
 ORIGIN
 Query Match 99.6%; Score 1389; DB 8; Length 3343;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CCATGCTGCTCTATGAAAAAGATGACATGCTGCTATATCCGTTTCTTAGGGTCC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGACATGCTGCTATATCCGTTTCTTAGGGTCC 60
 QY 61 CTTCTTTCGCTTATTAAGTCACTGAACTGGGGTTCAAAAACTTCCACGGGTGACATGAT 120
 DB 61 CTTCTTTCGCTTATTAAGTCACTGAACTGGGGTTCAAAAACTTCCACGGGTGACATGAT 120
 QY 121 CTCATGTTTCACTTCTCCACCTGCGTTGACATTTCTTGATGTCGCTGTTCCCAT 180
 DB 121 CTCATGTTTCACTTCTCCACCTGCGTTGACATTTCTTGATGTCGCTGTTCCCAT 180
 QY 181 CTGACCGAGGCCCATACGACACTTTCCGGAACCCATCAAGGCTTTCCGATGCCCA 240
 DB 181 CTGACCGAGGCCCATACGACACTTTCCGGAACCCATCAAGGCTTTCCGATGCCCA 240
 QY 241 CGAGAGGTATCGGGTGTGTGATCGAGGGATATATGTCCTCCCAATCGCATCTATA 300
 DB 241 CGAGAGGTATCGGGTGTGTGATCGAGGGATATATGTCCTCCCAATCGCATCTATA 300

Db 241 CGAGACGTATCGGGTCGTGGATCCAGGGATATATGTCCCCCAACATGTCACCTATA 300
Qy 301 TTATTTATCTTTAGATTTTATTTATTTTGGAAAAATACAAACTTATCTTTTGTGA 360
Db 301 TTATTTATCTTTAGATTTTATTTATTTTGGAAAAATACAAACTTATCTTTTGTGA 360
Qy 361 GGGCTCAGCATGATTTTGGCTTAAAGGGCCAGAAATGCGAGGACCGACCTGCTAGTG 420
Db 361 GGGCTCAGCATGATTTTGGCTTAAAGGGCCAGAAATGCGAGGACCGACCTGCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACCAAGTAATCTAATCCACT 480
Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACCAAGTAATCTAATCCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATTAATTAACCAACATCTCTTAAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATTAATTAACCAACATCTCTTAAAAAA 540
Qy 541 CAAGCATATTTCCAAAGAGACAATAATGTGTACAGTTTACAAACATCTAAGACGACAAA 600
Db 541 CAAGCATATTTCCAAAGAGACAATAATGTGTACAGTTTACAAACATCTAAGACGACAAA 600
Qy 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGTT 660
Db 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTATTTGTT 660
Qy 661 ATGTGTTTTATATACATTTCTCTCTTCAATAGTGATTTCTTCGATTTTAAATA 720
Db 661 ATGTGTTTTATATACATTTCTCTCTTCAATAGTGATTTCTTCGATTTTAAATA 720
Qy 721 ATGACTATAAAGTCATTTTATATTAAGAGACGATGCTAGATTTCTCGTCAAAAAATC 780
Db 721 ATGACTATAAAGTCATTTTATATTAAGAGACGATGCTAGATTTCTCGTCAAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCOAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGCTAGTTGGCAACCTGTTCTTTCOAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTATTTTCTCTTTTAAATAATGAAACACTTGAAGAAATAGAGT 900
Db 841 TTCAAAAAAATAGTATTTTCTCTTTTAAATAATGAAACACTTGAAGAAATAGAGT 900
Qy 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTATTAATTTTG 960
Db 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTATTAATTTTG 960
Qy 961 GCCAGCCCCCAATTAATTTTAAACCGAAATCGAAATGAGCGGAAACCAATCTGAGCTAT 1020
Db 961 GCCAGCCCCCAATTAATTTTAAACCGAAATCGAAATGAGCGGAAACCAATCTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
Db 1021 TTCTCTAGATTAAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
Qy 1081 TGAAGATGCGGGTTGGCAACGATAGCCACGTAATCATAGTCAATAGTGCTCTAGCTCA 1140
Db 1081 TGAAGATGCGGGTTGGCAACGATAGCCACGTAATCATAGTCAATAGTGCTCTAGCTCA 1140
Qy 1141 GGTTCGGAGACTCTCGTGTATCTCAGATGGAATCACTACATGCTTGTCAACGGTTCGTC 1200
Db 1141 GGTTCGGAGACTCTCGTGTATCTCAGATGGAATCACTACATGCTTGTCAACGGTTCGTC 1200
Qy 1201 TTGTTCCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATACCTAATCCCAACAT 1260
Db 1201 TTGTTCCATCGTCCAGACCTTGCTTATTTGAAACCAAGAGATACCTAATCCCAACAT 1260
Qy 1261 CCATCTTACTACATGCAATCTTCATGCAACACGCAATATGTTTCCGAAACCAATCATTT 1320
Db 1261 CCATCTTACTACATGCAATCTTCATGCAACACGCAATATGTTTCCGAAACCAATCATTT 1320
Qy 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
Db 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380

Qy 1381 CGTCCACCA 1389
Db 1381 CGTCCACCA 1389

RESULT 4
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoidae; Andropogoneae; Zea.

REFERENCE
1
AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source
1. 1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGCTGCTCTATGAAAAAGATGATCAATGCTGCTATTCGTTTCTTAAAGGCTCC 60
Db 1 CCATGCTGCTCTATGAAAAAGATGATCAATGCTGCTATTCGTTTCTTAAAGGCTCC 60
Qy 61 CTTCCTTGCCCTTATTAATGATGATGAGTGGGTTCACAAAACTTCCAGGGGTGATGAT 120
Db 61 CTTCCTTGCCCTTATTAATGATGATGAGTGGGTTCACAAAACTTCCAGGGGTGATGAT 120
Qy 121 CTCGATGTTCCACTTCTCCACCTCGCGTTGACATTTCTTGATGTCGGTGGTTCCCAT 180
Db 121 CTCGATGTTCCACTTCTCCACCTCGCGTTGACATTTCTTGATGTCGGTGGTTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAACACCTTTGCGGACACCCATCAAGGGCTTTCCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAACACCTTTGCGGACACCCATCAAGGGCTTTCCGATGGCCCA 240
Qy 241 CGAAGCGTATCGGGTCGTGGTATCCAGGGGATATATGTCGCCCAATCGTCACCTATA 300
Db 241 CGAAGCGTATCGGGTCGTGGTATCCAGGGGATATATGTCGCCCAATCGTCACCTATA 300
Qy 301 TTATTAATCTTTAGATATTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Db 301 TTATTAATCTTTAGATATTTATTTTAAATTTTGGAAAAATTAACAACTTATCTTTGTGTA 360
Qy 361 GGGCTCAGCATGATTTTGGCTTAAAGGGCCAGAAATGCGAGGACCGACCTGCTAGTG 420
Db 361 GGGCTCAGCATGATTTTGGCTTAAAGGGCCAGAAATGCGAGGACCGACCTGCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACCAAGTAATCTAATCCACT 480
Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACCAAGTAATCTAATCCACT 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATTAATTAACCAACATCTCTTAAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAATCATTAATTAACCAACATCTCTTAAAAAA 540
Qy 541 CAAGCATATTTCCAAAGAGACAATAATGTGTACAGTTTACAAACATCTAAGACGACAAA 600
Db 541 CAAGCATATTTCCAAAGAGACAATAATGTGTACAGTTTACAAACATCTAAGACGACAAA 600

QY 601 TTATATGAAAGGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTATTTGTT 660
DB 601 TTATATGAAAGGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTGTATTTGTT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
QY 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGTCGTAGTTCTGTTCAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGTCGTAGTTCTGTTCAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAAAGATTTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAAAGATTTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTTCTCTTATATAAATAGAAAAACATTAGAAAAATAGAT 900
DB 841 TTCAAAAAAATTAAGTTATTTTCTCTTATATAAATAGAAAAACATTAGAAAAATAGAT 900
QY 901 TGGCAGACTAGCCCTAGATGTTTTCCCAATTAATCAATCACTGTATTAATTTTGG 960
DB 901 TGGCAGACTAGCCCTAGATGTTTTCCCAATTAATCAATCACTGTATTAATTTTGG 960
QY 961 GCGAGCCCAATTAATTTTAAACCGAACTGAATGAGCGAAACCAATTCGAGCTAT 1020
DB 961 GCGAGCCCAATTAATTTTAAACCGAACTGAATGAGCGAAACCAATTCGAGCTAT 1020
QY 1021 TTCTAGATTTAGTATAAAGGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
DB 1021 TTCTAGATTTAGTATAAAGGAGAGAGAGAGAAATCAGTTTAACTGATTTGCC 1080
QY 1081 TGAAGTGTGGGTTTGGCAACGATAGCCGATTAATGCTATAGTGTGCTTACGTC 1140
DB 1081 TGAAGTGTGGGTTTGGCAACGATAGCCGATTAATGCTATAGTGTGCTTACGTC 1140
QY 1141 GGTTCGGAGCTCTCGGTGATCTCAATAGGCAATACGTTGTTCAACCGTTGTC 1200
DB 1141 GGTTCGGAGCTCTCGGTGATCTCAATAGGCAATACGTTGTTCAACCGTTGTC 1200
QY 1201 TTGTTCCATGCTCAAGCCTTGCTTATTCGAAACCAAGAGATACCTACCCAAACAT 1260
DB 1201 TTGTTCCATGCTCAAGCCTTGCTTATTCGAAACCAAGAGATACCTACCCAAACAT 1260
QY 1261 CCATCTTACTACTGCAACCTTCATGCAACGCAATGTTTCTGTAAC 1311
DB 1261 CCATCTTACTACTGCAACCTTCATGCAACGCAATGTTTCTGTAAC 1311

RESULT 5
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177
ACCESSION BD062177.1 GI:22607782
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and
Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
PIONEER HI BRED INTERNATIONAL INC
JP 2001520523-A/2
COMMENT
BD 30-OCT-2001 JP 199504910
PR 19-JUN-1998 JP 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
TIMMY L KENDALL

PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key location/Qualifiers.
FEATURES
source
1.1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 94.0%; Score 1311; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGTGTCTATGAAAAAGATGATGTAATGCTATATCGTTTCTTAGGGTCC 60
DB 1 CCATGGTGTCTATGAAAAAGATGATGTAATGCTATATCGTTTCTTAGGGTCC 60
QY 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTACAAAAAATTCCAGGGTGATGAT 120
DB 61 CTTCTTCTGCTTATTAAGTGAATCGGGTTACAAAAAATTCCAGGGTGATGAT 120
QY 121 CTCGATGTTCCACTTTCCTCCCACTCGCGTTGCAATTTCTTGAGATCGGTGCTCCAT 180
DB 121 CTCGATGTTCCACTTTCCTCCCACTCGCGTTGCAATTTCTTGAGATCGGTGCTCCAT 180
QY 181 CTGACCGAGGCCCATGACGACCTTTCCGAGCAACCCATCAAGGGCTTTTCGATGGCCCA 240
DB 181 CTGACCGAGGCCCATGACGACCTTTCCGAGCAACCCATCAAGGGCTTTTCGATGGCCCA 240
QY 241 CGAGAGCTATCGGGTGTGATGTCAGGGGATATATGTCGCCCAATCGTCACTATA 300
DB 241 CGAGAGCTATCGGGTGTGATGTCAGGGGATATATGTCGCCCAATCGTCACTATA 300
QY 301 TTAATATCTTTAGATATTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
DB 301 TTAATATCTTTAGATATTTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
QY 361 GGGCTCAGATGATTTTTCGCTTAGGGCCAGAAATGCGAGGACCGACATGCTAGTG 420
DB 361 GGGCTCAGATGATTTTTCGCTTAGGGCCAGAAATGCGAGGACCGACATGCTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATACCAAGTAACTATCCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAAAGATTTAAAAAATACCAAGTAACTATCCACT 480
QY 481 CGAAAGCTATCATGTATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAA 540
DB 481 CGAAAGCTATCATGTATATGTTTAAAGAAACATCTATTAACACGATCCTTTAAAAA 540
QY 541 CAAGCATATTTGAAAGAGACAAATATGTTTCAAGTTTCAACATCTAAGAGCGACAA 600
DB 541 CAAGCATATTTGAAAGAGACAAATATGTTTCAAGTTTCAACATCTAAGAGCGACAA 600
QY 601 TTATATCGAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATTTGTT 660
DB 601 TTATATCGAAGGTAAGCTATGACGTTCAAGATTTTCTTTTCAATCTGTATTTGTT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTTACAAATAGATGATTTTCTCCGATTTTATATA 720
QY 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGTCGTAGTTCTGTTCAAAATC 780
DB 721 ATGACTATAAAGTCATTTTATATATAGAGACGCGATGTCGTAGTTCTGTTCAAAATC 780
QY 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAAAGATTTTGATTTT 840
DB 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCCGTTTCTTCAAAAGATTTTGATTTT 840

OY	841	TTCAAATAAAATTAAGTTATTTCTCTTATTAATAATAGAAAACATTGAAAAAATAGAGT	900
Db	841	TTCAAAAAAATTAAGTTATTTCTCTTATTAATAATAGAAAACATTGAAAAAATAGAGT	900
OY	901	TGCCAGACTAGCCCTAGATGTTTTCCCAATTAATTAACATCAGCTGATATATATTTTG	960
Db	901	TGCCAGACTAGCCCTAGATGTTTTCCCAATTAATTAACATCAGCTGATATATATTTTG	960
OY	961	GCCAGCCCCATTAATTAATTTAAACCGAACTGAATTCGAGCGAAACCAATTTGAGCTAT	1022
Db	961	GCCAGCCCCATTAATTAATTTAAACCGAACTGAATTCGAGCGAAACCAATTTGAGCTAT	1022
OY	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAGTCATGTGCC	1080
Db	1021	TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAGTCATGTGTCCC	1080
OY	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCAATAGCTCATAGGTCGACGTCA	1140
Db	1081	TGAGATGTGCGGTTTGGCAACGATAGCCACCGTAATCAATAGCTCATAGGTCGACGTCA	1140
OY	1141	GGTTGCGAGGCTCTGCTGTCAATCTCAATGGATATCTACAGCTGTTCACACCGTTGCTC	1200
Db	1141	GGTTGCGAGGCTCTGCTGTCAATCTCAATGGATATCTACAGCTGTTCACACCGTTGCTC	1200
OY	1201	TTGTTTCATCGTCCAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAAT	1266
Db	1201	TTGTTTCATCGTCCAAGCCTTGCTATTCTGAACCAAGAGATACCTATCCCAACAAT	1266
OY	1261	CCATTTTACTCATGCACTTCATGCAAAACAGCACATATGTTTTCTGTAC	1311
Db	1261	CCATTTTACTCATGCACTTCATGCAAAACAGCACATATGTTTTCTGTAC	1311

LOCUS	AX224402	255 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 9 from Patent WO0160997.				
ACCESSION	AX224402				
VERSION	AX224402.1 GI:15554644				
KEYWORDS	.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
AUTHORS	1 Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.				
TITLE	Male tissue-prefered regulatory region and method of using same				
JOURNAL	Patent: WO 0160997-A 9 23-AUG-2001;				
FEATURES	PIONEER HI-BRED INTERNATIONAL, INC. (US)				
SOURCE	Location/Qualifiers				
	1..255				
	/Organism="Zea mays"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4577"				
ORIGIN					
Query Match:	11.3%; Score 157; DB 6; Length 255;				
Best Local Similarity	100.0%; Pred. No. 4.9e-71;				
Matches 157; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
Oy	1155	CGTGCATCTCACACGAGCATACTACATGCTGTTCAACGGTTCGCTTGCATCGTCC	1214		
Db	15	CGTGCATCTCACATGAGCATACTACATGCTGTTCAACGGTTCGCTTGCATCGTCC	74		
Oy	1215	AAGCCTTGCCATTCTGGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG	1274		
Db	75	AAGCCTTGCCATTCTGGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG	134		
Oy	1275	CAACTTCATGCAACACGACATATGTTTCTCTGAC	1311		
Db	135	CAACTTCATGCAACACGACATATGTTTCTCTGAC	171		

	RESULT 7			
	AX224396			
LOCUS	AX224396	158 bp	DNA	linear
DEFINITION	Sequence 3 from Patent WO0160997.			PAT 10-SEP-2001
ACCESSION	AX224396			
VERSION	AX224396.1	GI:15554638		
KEYWORDS	.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
REFERENCE	1			
AUTHORS	Albertsen, M.C., Fox, T.W., Garmaat, C.W., Huffman, G. and Kendall, T.L.			
TITLE	Male tissue-preferred regulatory region and method of using same			
JOURNAL	Patent: WO 0160997-A 3 23-Aug-2001;			
	PIONEER HI-BRED INTERNATIONAL, INC. (US)			
FEATURES	location/Qualifiers			
source	1..158			
	/organism="Zea mays"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:4577"			
ORIGIN				
	Query Match	8.0%; Score 111; DB 6; Length 158;		
	Best Local Similarity	100.0%; Pred. No. 1e+46;		
	Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
CY	1201 TTGTTCCATCGTCCAAGCCTTGCTATTTCGAACCAAGAGATCTACTATCCCAAACAAT	1260		
DB	48 TTGTTCATCGTCCAAGCCTTGCTATTTCGAACCAAGAGATCTACTATCCCAAACAAT	107		

AX224398	LOCUS	AX224398	50 bp	DNA	linear	PAT 10-SEP-2001
DEFINITION	Sequence 5 from Patent WO0160997.					
ACCESSION	AX224398					
VERSION	AX224398.1	GI:15554640				
KEYWORDS						
SOURCE						
ORGANISM	Zea mays					
	Zea mays					
	Eumariyote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
REFERENCE	1					
AUTHORS	Albersen, M. C., Fox, T. W., Garnaat, C. W., Hiffman, G. and Kendall, T. L.					
TITLE	Male tissue-preferred regulatory region and method of using same					
JOURNAL	Patent: WO 0160997-A 5 23-AUG-2001					
FEATURES	PIONEER	HI-BRED INTERNATIONAL, INC. (US)				
	location/Qualifiers					
source	1..50					
	/organism="Zea mays"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:4577"					
ORIGIN						
Query Match	3.6%; Score 50; DB 6; Length 50;					
Best Local Similarity	100.0%; Pred. No. 1.8e-14;					
Matches	50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1229	CTGAACCAAGGATACCTACTCCCAACAACATCCATCTTACTCATGCAC	1278			
Db	1	CTGAACCAAGGATACCTACTCCCAACAACATCCATCTTACTCATGCAC	50			
RESULT	9					
	AX224399					

LOCUS AX224399 40 bp DNA linear PAT 10-SEP-2001
 DEFINITION Sequence 6 from Patent WO0160997.
 ACCESSION AX224399 GI:15554641
 VERSION
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1
 AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 0160997-A 6 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source 1..40
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"
 ORIGIN
 Query Match 2.9%; Score 40; DB 6; Length 40;
 Best Local Similarity 100.0%; Pred. No. 3.4e-09;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1239 AGGATACCTACTCTCCCAACATCTTACTCATGCAAC 1278
 |||||||
 1 AGGATACCTACTCTCCCAACATCTTACTCATGCAAC 40
 |||||||
 RESULT 10
 AX224397 30 bp DNA linear PAT 10-SEP-2001
 LOCUS AX224397
 DEFINITION Sequence 4 from Patent WO0160997.
 ACCESSION AX224397
 VERSION AX224397.1 GI:15554639
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1
 AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 0160997-A 4 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source 1..30
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"
 ORIGIN
 Query Match 2.2%; Score 30; DB 6; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.00067;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1179 CATGCTTTTCACCGTCTGCTTTGTTCCA 1208
 |||||||
 1 CATGCTTTTCACCGTCTGCTTTGTTCCA 30
 |||||||
 RESULT 11
 AC149836 114817 bp DNA linear HTG 22-JUN-2004
 LOCUS AC149836
 DEFINITION Zea mays clone ZMMB0496L17, *** SEQUENCING IN PROGRESS ***, 3
 unoriented pieces.
 ACCESSION AC149836
 VERSION AC149836.1 GI:49035067
 KEYWORDS HTG; HTGS_PHASE1; HTGS_FULFLOP; HTGS_ACTIVEFIN.
 SOURCE Zea mays

ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 114817)
 AUTHORS Birren, B., Nuebaum, C., Lander, E., Butler, E., Wang, R., Bharti, A.K. and Messing, J.
 TITLE Zea mays, clone ZMMB0496L17
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 114817)
 REFERENCE 1
 AUTHORS Birren, B., Nuebaum, C., Lander, E., Butler, E., Wang, R., Bharti, A.K., Messing, J., Abouelkheil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H.M., Barina, N., Baeten, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collimore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horvath, J., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Minova, T., Miens, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, V., Raymond, C., Retta, R., Rice, C., Rogov, P., Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talmes, J., Tefaye, S., Theodore, J., Toppan, K., Travers, M., Vasilev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 DIRECT SUBMISSION
 Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@broad.mit.edu

 Bharti, A.K. and Messing, J.: The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers, The State University of New Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA (http://pgr.rutgers.edu)
 Butler, E. and Wang, R.: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
 ----- Project Information
 Center project name: L30352
 Center clone name: 496_L_17

 * NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 35156: contig of 35156 bp in length
 * 35157 35256: gap of 100 bp
 * 35257 97940: contig of 62684 bp in length
 * 97941 98040: gap of 100 bp
 * 98041 114817: contig of 16777 bp in length.
 Location/Qualifiers
 1..114817
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /culivar="B73"
 /db_xref="taxon:4577"
 FEATURES
 source

ORIGIN /clone="ZMMBBC0496L17" /clone_lib="CHORI-201 Maize B73 BAC Mbol Library"

Query Match 1.8%; Score 25; DB 2; Length 114817;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 844 AAAAAAAAAATGATTTCTCTCTT 868
Db 27886 AAAAAAAAAATGATTTCTCTT 27910

RESULT 12
AE010607
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source

AE010607 10384 bp DNA linear BCT 25-MAR-2002
Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 149 of
197 of the complete genome.
AE010607 AE009951
AE010607.1 GI:19714578

Fusobacterium nucleatum subsp. nucleatum ATCC 25586
Fusobacterium nucleatum subsp. nucleatum ATCC 25586
Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
Fusobacterium.
1 (bases 1 to 10384)
Kaprali,V., Anderson,I., Ivanova,N., Reznik,G., Los,T.,
Lykidis,A., Bhattacharyya,A., Bartman,A., Gardner,W., Grechkin,G.,
Zhu,L., Vasileva,O., Chu,L., Kogan,Y., Chaga,O., Goldsman,E.,
Bernal,A., Larsen,N., D'Souza,M., Walunas,T., Pusch,G.,
Hasselkorn,R., Fontein,M., Kyripides,N. and Overbeek,R.
Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586
J. Bacteriol. 184 (7), 2005-2018 (2002)

21886394
11889109
2 (bases 1 to 10384)
Kaprali,V., Anderson,I., Ivanova,N., Reznik,G., Los,T.,
Lykidis,A., Bhattacharyya,A., Bartman,A., Gardner,W., Grechkin,G.,
Zhu,L., Chu,L., Kogan,Y., Chaga,O., Goldsman,E., Bernal,A.,
Larsen,N., D'Souza,M., Walunas,T., Pusch,G.D., Hasselkorn,R.,
Fontein,M., Kyripides,N. and Overbeek,R.
Direct Submission
Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park
Drive, Chicago, IL 60612, USA
Location/Qualifiers
1. .10384
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC
25586"
/mol_type="genomic DNA"
/strain="ATCC 25586"
/sub_species="nucleatum"
/db_xref="ATCC:25586"
/db_xref="taxon:190304"
162. .1664
/gene="FN0998"
162. .1664
/gene="FN0998"
162. .1664
/gene="FN0998"
/transl_table=11
/product="Dipeptide-binding protein"
/protein_id="AAL95194.1"
/db_xref="GI:19714579"
/translation="MKKIFYLPLFVLSILVGGCGKNTSGPNSGVVIGGAKPKSIDP
HMYNSPDLVSRQFVNTLPSRKDGSIRPELAESYRKNDKLDIILKKGAKFHDGS
ELTADVVSFERMKDKPGSSIVESIDVEKVDYEIKLPLKSSSALLVNLHPT
SIVNKRYEAGNDLSVAPMGTAFLIAYNDGEKILEAFKQYFEGAPVKEKITPFI
PEDTSMALLETGEVDIATGMPVSTQITLSEANDKLDLISEPTATYEICINYEKTFAD
NKDFKALNVAIDKQSIIDISIFSGRGVAKSIVPNVFGYDGFEEYFPNPEAKELI
EKSGVMSFSLVYNDSPVRLVAQIIONLKHVGIDMNIETLEMGTYLOKTDGSGFS
AYIGMISGTSDDIVLYPLDLSKSTIGFQNRNRRYRNPEDKVEKARIALKPEERKE
HFKNAGIISQDSPLIVLYKNENIGINKRIKGFEDPTTMKFKNLEIK"
1669. .2712

gene
CDS

CDS
/gene="FN0999"
1669. 2712
/gene="FN0999"
/EC_number="3.4.11.-"
/codon_start=1
/transl_table=11
/product="Deblocking aminopeptidase"
/protein_id="AAL95195.1"
/db_xref="GI:19714580"
/translation="MILNRIDKTYLAKTVELLIRPSVGYTHNAIEVYQELKGLGI
KNYITRKGAALVYIKGDSNYKMI SHAVDVLGVVVKIKKSARLEVTNNGGFAMGS
VEGENVTIHTISGTYSGTLPLIKASVHVGVDAEMPRTERTEIRIDEDKTEDV
LKLGIGDFFISFETHTIRILDNGYKSRVLDKLCVAQILSYIKLKNKMLKPKDLY
IYFSNVEIRIGVSVFPEDLDEFLIYVDIGLVAGDEADGKVOJIIAKOSRETYFTL
RKLOETSKNNIKYTVGVNRYSGDATVAVLGDFPYACIGPVNDAVTHHERCHND
GITEVTKLIAYL"
2854. .3936
/gene="FN1000"
2854. .3936
/gene="FN1000"
/EC_number="2.8.1.6"
/codon_start=1
/transl_table=11
/product="Biotin synthase"
/protein_id="AAL95196.1"
/db_xref="GI:19714581"
/translation="WLKERNSAGGKFNFPFLKEKNNNOAEPNSVKEFISYLNKII
NEKYITREBAIFLSQIPNNMETNIIIPNADOIRBVPCCGYPLCTIIANSKSRS
ENCKTCAOSVHKGTIDVYGLISKELVYKARNENBAHRPSLTSGLGNGKEL
DKLVIVIKYIEHNSKELCASHGICTEALQKVDACVLYVYHNHLESSRRYPVCT
SHSYDNRNITIKNAVGLDVCSGIGFGLGTEIDRLMALDLRALEISYVDIVLVT
IPGTFENNEENVPPEILKTIISYIFMPKFLRYCGGRIRKIGBAKTKRGINSAL
TGNFLTGTITETDKNMKEGYEI"
3926. .4585
/gene="FN1001"
3926. .4585
/gene="FN1001"
/EC_number="6.3.3.3"
/codon_start=1
/transl_table=11
/product="Dechthiobiotin synthetase"
/protein_id="AAL95197.1"
/db_xref="GI:19714582"
/translation="MKFDPFVIGDITDVGKTYVSTLYKALKRYNFOYKXIOSGCF
LKDGKLTAPDVDFLTKFVGDYDQSMVYTLKEVSPHASEMEGIRIEIENKHE
DLKKGHNIIVBAGAGLYVPLIRDFIYIYDLIKLFNLVNVVYCCGYKGNATNTMLTLN
ALDTWGRKDHGLVFNRYRGQFFEDNIRVALLSKIKNYLIILKNGQKEISDKTEKFP
N"
4653. .6011
/gene="FN1002"
4653. .6011
/gene="FN1002"
/EC_number="2.6.1.62"
/codon_start=1
/transl_table=11
/product="Adenosylmethionine-8-amino-7-oxononanoate
aminotransferase"
/protein_id="AAL95198.1"
/db_xref="GI:19714583"
/translation="WRGVCTIMINLSSEIQKDKLVYFHPQCMQKDPENPPLVYKGD
GLYLIDENGNKYMDCISSWVNLFGHCNKRINRIITTEOVNMLEHVI FANFTHEPAEL
CEELTKVLPKGINKEFLPSDNGSSCIEMALKSPYHLOQGNPOKTFSLBNAYHGET
IGALGVGDVIDFETERYPLIKGRKRVYIDSKLSDPEFVLEDECIKELDLIILKN
HNETACIMVEPMVOGAARIKIYSAKEFLAARDLTKVNIHLIDDSIIMGFGRTGMPA
CEHAIEEDPMACIKAGSSGYPIAMLCITTDIFNAFADYKGRSFLHSHTYSGNPL
GGRILAEVLRI FKEENIILKTINEKGYIRANMKEIFEDKSYIKQIRNLGILGAILDK
NLPVNRIGKEIYFALKKGAFLRIGNSVYFMPYVITYEIDMKLHVCKESIBEL
KI"
6737. .7558
/gene="FN1003"
6737. 7558
/gene="FN1003"
/codon_start=1

gene
CDS

RESULT	13
LOCUS	AC113237/c
DEFINITION	Canis familiaris clone RP81-70J12, WORKING DRAFT SEQUENCE, 3 ordered pieces.
ACCESSION	AC113237
VERSION	AC113237.2 GI:22128672
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE	Canis familiaris (dog)
ORGANISM	Canis familiaris
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
AUTHORS	1 (bases 1 to 184509) Ahter,N., Antcnellos,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Grantle,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P., Lee-Jin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B., Margulis,E.H., Mastello,C., Maekel,B., Maurrian,S.D., McElroskey,J.C., McDowell,J., Pegunigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Schueler,M.G., StancilPop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Tsurgoon,C., Vogt,J.L., Walker,M.A., Weibachy,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D. NISC Comparative Sequencing Initiative Unpublished 2 (bases 1 to 184509) Green,E.D. Direct Submission Submitted (27-FEB-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA 3 (bases 1 to 184509) Green,E.D. Direct Submission Submitted (07-AUG-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA On Aug 7, 2002 this sequence version replaced gi:18958664.
TITLE	Genome Center
JOURNAL	Center: NIH Intramural Sequencing Center
REFERENCE	Center code: NISC
AUTHORS	Web site: http://www.nisc.nih.gov
TITLE	Contact: nisc.zoo@hgri.nih.gov
JOURNAL	Project Information
COMMENT	Center project name: axv Center clone name: 070J12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.990319
Consensus quality: 184204 bases at least Q40

Insert size: 162000; agarose-fp
Insert size: 184309; sum-of-contigs
Quality coverage: 10.51x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs


```

* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced.
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 127823: contig of 127823 bp in length
127824 127923: gap of unknown length
127924 147441: contig of 19518 bp in length
147442 147541: gap of unknown length
147542 184509: contig of 36968 bp in length.
FEATURES
Source
1. 184509
/organism="Canis familiaris"
/mol_type="genomic DNA"
/db_xref="taxon:9615"
/cclone="RP81-70312"
/cclone_11b="RP81"
1. 127823
/clone="assembly_fragment"
/clone_end:T7
vector_side:left"
misc_feature
1. 3358
/note="clone overlaps with GenBank Accession Number
AC110671 clone RP81-69D24 (center project name aw)"
127924. 147441
/note="assembly_fragment"
129971. 184509
/note="clone overlaps with GenBank Accession Number
AC113567 clone RP81-265M12 (center project name cio)"
147542. 184509
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
ORIGIN
Query Match 1.7%; Score 24; DB 2; Length 184509;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1036 AAAAGGAGAGAGAGAGAGAGAA 1059
DB 23292 AAAAGGAGAGAGAGAGAGAA 23269
RESULT 14
AC098210/c 212621 bp DNA linear HTG 13-MAY-2003
LOCUS AC098210
DEFINITION Rattus norvegicus clone CH230-176E2, WORKING DRAFT SEQUENCE.
ACCESSION AC098210.8 GI:30581396
VERSION HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 212621)
Muzny,D,Marle,M, Metzker,M, Lee, A, Abramson, S., Adams C., Alder, J.,
Allen, C., Allen, H., Aisbrook, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavaros, I., Cessari, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Ditya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duvall, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gabriel, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebreyorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

```

```

Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
Hollins, B., Howells, S., Hulyk, S., Hune, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyet, A.,
Karcachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshew, L., Loulaged, H., Lozano, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundiada, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Parks, K.,
Nwackemehe, O., Okunodu, G., Olampunasegon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.,
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajic, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabox, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umanat, K.,
Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wiczysk, R., Wooden, H., Woley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.
Direct Submission
Unpublished
2 (bases 1 to 212621)
Worley, K. C.
Submitted (13-OCR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 212621)
Rat Genome Sequencing Consortium.
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced gi:23268092.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
----- Project Information
Contact: hgsc-help@bcm.tmc.edu
Center project name: GILN
Center clone name: CH230-176E2
----- Summary Statistics
Assembly program: ATLAS 3.0;
Consensus quality: 205647 bases at least Q40
Consensus quality: 207827 bases at least Q30
Consensus quality: 209298 bases at least Q20

```


[illegible]

Kwiatkowski, C.L., Lebowitz, H., Lewy, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
 Lorenz, L., Louis, S., Lozano, R., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmood, M., Mallory, K., Mangum, A.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,
 Miosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokwelen, O., Okunolu, G., Olariu, S., Pal, S., Parks, K.,
 Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
 Plopper, F., Polidexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richard, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
 Sanders, W., Savary, G., Scheier, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smail, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,
 Steinle, M., Strong, R., Sutton, A., Swack, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, E., Usmani, K.,
 Valas, R., Vera, V., Villaseña, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wlezyk, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhauser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstein, G., and Gibbs, R.A.
 Rat Genome Sequencing Consortium.
 Direct Submission
 Unpublished
 2 (bases 1 to 213753)
 Direct Submission
 Morley, K.C.
 Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 213753)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:22698201.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSTM
 Center clone name: CH230-155J23
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 203043 bases at least Q40
 Consensus quality: 204685 bases at least Q30
 Consensus quality: 205672 bases at least Q20
 Estimated insert size: 206221; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 212625: contig of 212625 bp in length
* 212626 212725: gap of unknown length
* 212726 213753: contig of 1028 bp in length.
Location/Qualifiers

FEATURES
source

1. .213753
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-155023"
1. .1535
misc_feature
/note="wgs end_extension
clone_end:Sp6"
1586. .2703
misc_feature
/note="wgs end_extension
clone_end:Sp6"
5028. .5742
misc_feature
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BH364626"
33891. .35390
misc_feature
/note="wgs_contig"
211446. .212625
misc_feature
/note="wgs_contig"

ORIGIN

Query Match 1.7%; Score 24; DB 2; Length 213753;
Best Local Similarity 100.0%; Pred.No. 1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATGATTATTTCTTTAT 871
|||||
Db 65390 AAAATGATTATTTCTTTAT 65367

Search completed: September 16, 2005, 15:25:37
Job time : 4110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 08:08:50 ; Search time 578.5 Seconds

(without alignments)
14264.684 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccatggtgtctctcatgaataa.....cttttcgtccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: genebegn1980s:*
3: genebegn1980s:*
4: genebegn2000s:*
5: genebegn2001as:*
6: genebegn2002as:*
7: genebegn2002bs:*
8: genebegn2003as:*
9: genebegn2003bs:*
10: genebegn2003cs:*
11: genebegn2003ds:*
12: genebegn2004as:*
13: genebegn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	2	AAK07408 Zea mays
2	1394	100.0	1394	5	AAH76332 Z. mays M
3	1311	94.0	1394	2	AAK07409 Zea mays
4	1311	94.0	1394	5	AAH76333 Z. mays M
5	157	11.3	8.0	5	AAH76340 Z. mays M
6	111	8.0	158	5	AAH76334 Z. mays M
7	50	3.6	50	5	AAH76336 Z. mays M
8	40	2.9	40	5	AAH76337 Z. mays M
9	30	2.2	30	5	AAH76335 Z. mays M
10	23	1.6	849	6	ABL40437 Sunflower
11	21	1.5	333	5	ABV56505 Human pro
12	21	1.5	459	4	ABAS8867 Human foe
13	21	1.5	459	4	AAI38581 Probe #72
14	21	1.5	459	4	AAK37768 Human bon
15	21	1.5	459	4	AAK07027 Human bra
16	21	1.5	459	4	ABK32487 Human liv
17	21	1.5	459	4	ABK07566 Human gen
18	21	1.5	965	6	AAAB69660 Pinus rad
19	21	1.5	9255	6	ABL34152 Human imm
20	21	1.5	9265	6	AAK63354 Chemical1

21	21	1.5	18817	6	ABL34494 Human met
22	21	1.5	18817	6	ABL70161 Chemical1
23	21	1.5	18817	7	ADK99755 Bisluphic
24	21	1.5	26493	12	ADJ12386 DNA fragm
25	20	1.4	260	5	ABV60726 Human pro
26	20	1.4	680	6	ABO18134 Oligonuc1
27	20	1.4	680	6	ABO18135 Oligonuc1
28	20	1.4	2523	2	AAK76376 Zea mays
29	20	1.4	10439	6	AAK47817 Alternati
30	20	1.4	11779	6	AAK42891 Maize sug
31	20	1.4	51198	11	ACN45000 Mouse gen
32	20	1.4	91071	11	ACN44004 Mouse gen
33	20	1.4	110000	6	ABA90193_2 Continuation (3 of
34	20	1.4	110000	6	ABA90193_3 Continuation (4 of
35	20	1.4	110000	6	ABQ87681_2 Continuation (3 of
36	20	1.4	110000	6	ABQ87681_3 Continuation (4 of
37	20	1.4	110000	8	ABX33717_2 Continuation (3 of
38	20	1.4	110000	8	ABX33717_3 Continuation (4 of
39	20	1.4	148497	12	ADQ18925 Human scf
40	20	1.4	194534	12	ADQ97481 Human can
41	20	1.4	337344	13	ABD32715 Human can
42	19	1.4	418	13	ADR65258 Cotton CD
43	19	1.4	468	2	AAK19065 Human PPA
44	19	1.4	485	4	ABA57528 Human foe
45	19	1.4	485	4	AAI37083 Probe #57

ALIGNMENTS

RESULT 1	AAK07408 standard; DNA; 1394 BP.
ID	AAK07408
AC	AAK07408;
XX	08-JUN-1999 (first entry)
DE	Zea mays M45 male tissue-preferred regulatory region.
XX	M45; male; tissue-preferred; regulatory region; plant cells;
KM	plant tissue; differentiated; maize; hybrid seed; fertility; ss.
OS	Zea mays.
XX	WO9859061-A1.
PN	30-DEC-1998.
PD	19-JUN-1998; 98WO-US012895.
XX	23-JUN-1997; 97US-00880499.
PR	(PION-) PIONEER HI-BRED INT INC.
PA	Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
PI	WPI; 1999-105628/09.
DR	New nucleic acid encoding a M45 male tissue-preferred regulatory region
PT	- useful in mediating plant fertility, especially hybrid seed production.
XX	Claim 2; Page 22-23; 39pp; English.
XX	The sequence is that encoding an M45 male tissue-preferred regulatory
CC	region. It may be used in the construction of a vector for a method of
CC	producing exogenous genes in a male tissue- preferred manner; which is
CC	useful in restoring or conferring fertility, such as in hybrid seed
CC	production. In conferring fertility, a monocot/dicot plant is transformed
CC	with the exogenous nucleotide sequence (a male sterility gene, preferably
CC	M45), which encodes a product selected from auxins, rolB and diptheria
CC	toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC	and infertile plants

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other:
 SQ Query Match 100.0%; Score 1394; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGGTGTCTCTATGAAAAAGATGAGTACATGTCATATCCGTTTCTTAGGGTCC 60
 Db 1 CCATGGTGTCTCTATGAAAAAGATGAGTACATGTCATATCCGTTTCTTAGGGTCC 60

Qy 61 CTTCCTTCGCTTATTACTGAATCGGGTTTACAAAACTTCCACGGGTGATGAT 120
 Db 61 CTTCCTTCGCTTATTACTGAATCGGGTTTACAAAACTTCCACGGGTGATGAT 120

Qy 121 CTCCATGTTCCACTTCCCACTCGGGTTCACATTTCTTGATGTCGGTGGTCCAT 180
 Db 121 CTCCATGTTCCACTTCCCACTCGGGTTCACATTTCTTGATGTCGGTGGTCCAT 180

Qy 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCTTTGAGTGGCCCA 240
 Db 181 CTGACCGAGGCCCATCGACACCTTTGGGACACCCATCAAGGGCTTTGAGTGGCCCA 240

Qy 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCTCCCAATGTCACCTATA 300
 Db 241 CGAAGCGTATCGGGTCGTGTGATCGAGGGATATATGTCCTCCCAATGTCACCTATA 300

Qy 301 TTATATCTCTTGAATTTTAAATTTTGAAGAAAAATCAAACTTATCTTTGGTGA 360
 Db 301 TTATATCTCTTGAATTTTAAATTTTGAAGAAAAATCAAACTTATCTTTGGTGA 360

Qy 361 GGGGCTCAGCATAGATTTGCTTGAAGGCCAGAAATGAGAGACCGCATGCTAGTG 420
 Db 361 GGGGCTCAGCATAGATTTGCTTGAAGGCCAGAAATGAGAGACCGCATGCTAGTG 420

Qy 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480
 Db 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACCAAGTAACTATCCACT 480

Qy 481 CGAAGCTATCAGTATATGTTTAAAGAACATCATTTAAACCAACATCCTCTTAAAAA 540
 Db 481 CGAAGCTATCAGTATATGTTTAAAGAACATCATTTAAACCAACATCCTCTTAAAAA 540

Qy 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600
 Db 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTACAAACATCTAAGACGACAA 600

Qy 601 TTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTTATTTGTT 660
 Db 601 TTATATGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCAATCTTGTTATTTGTT 660

Qy 661 ATTGTTTTTATATACATTTCTTCTTACATATGAGTATTTTCTTCGATTTTAAAA 720
 Db 661 ATTGTTTTTATATACATTTCTTCTTACATATGAGTATTTTCTTCGATTTTAAAA 720

Qy 721 ATGACTATAAAGTCATTTTATATTAAGAGCAGCATGTCGTAGATTTCTCGTTCAAAAATC 780
 Db 721 ATGACTATAAAGTCATTTTATTAAGAGCAGCATGTCGTAGATTTCTCGTTCAAAAATC 780

Qy 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTTGATTTT 840
 Db 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTCAAGAAATTTTGATTTT 840

Qy 841 TTCAAAAAAATAGTTATTTTCTCTTATAAAAATGAAAAACATAGAAAAATGAGT 900
 Db 841 TTCAAAAAAATAGTTATTTTCTCTTATAAAAATGAAAAACATAGAAAAATGAGT 900

Qy 901 TGCAGAGTACCCCTAGATGTTTTCCCAATAAATTAACAATCACTGTATTAATTTTG 960
 Db 901 TGCAGAGTACCCCTAGATGTTTTCCCAATAAATTAACAATCACTGTATTAATTTTG 960

Qy 961 GCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCTAT 1020
 Db 961 GCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCTAT 1020

Db 961 GCCAGCCCATTAATTTTAAACCGAACTGAAATGAGCGAAACCAATCTGAGCTAT 1020
 Qy 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCTGTTTAAGTCAATGTCCTC 1080
 Db 1021 TTCTCTAGATTTAGTAAAAAGGAGAGAGAGAGAAATCTGTTTAAGTCAATGTCCTC 1080

Qy 1081 TGAGATGTCGGGTTTGGCAACATAGACCCGTAATCATAGCTCATAGGTGCTACGTCA 1140
 Db 1081 TGAGATGTCGGGTTTGGCAACATAGACCCGTAATCATAGCTCATAGGTGCTACGTCA 1140

Qy 1141 GGTTCGGAGCTCTCGTGTCAATGCAATGAGATACATGATCTGTTCAACGGTTCGTC 1200
 Db 1141 GGTTCGGAGCTCTCGTGTCAATGCAATGAGATACATGATCTGTTCAACGGTTCGTC 1200

Qy 1201 TTGTTCCATGTCGCAAGCCTTCCATTTCTGAAACCAAGGATATCTACTCCAAACAT 1260
 Db 1201 TTGTTCCATGTCGCAAGCCTTCCATTTCTGAAACCAAGGATATCTACTCCAAACAT 1260

Qy 1261 CCATCTTACTCATGCACTTCCATGCAAAACAGCATATGTTTCTGAAACCAATCATT 1320
 Db 1261 CCATCTTACTCATGCACTTCCATGCAAAACAGCATATGTTTCTGAAACCAATCATT 1320

Qy 1321 AAAGATCACAACAGCTAGCGGTTCTCCGCTAGCTTCCCTCTCTGCGATCTTTT 1380
 Db 1321 AAAGATCACAACAGCTAGCGGTTCTCCGCTAGCTTCCCTCTCTGCGATCTTTT 1380

Qy 1381 CGTCCACACCATG 1394
 Db 1381 CGTCCACACCATG 1394

RESULT 2
 AAH76332
 ID AAH76332 standard; DNA; 1394 BP.
 XX
 AC AAH76332;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE 2. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 OS Zea mays.
 PN WO200160997-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 XX
 PI Albrechtsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 XX WPI; 2001-514772/56.
 DR
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 XX essential for initiating transcription of the Ms45 gene useful for
 XX mediating fertility in a male plant.
 PS Claim 4; Page 46; 50pp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (I)
 XX comprising nucleotide sequences essential for initiating transcription of
 XX the Ms45 gene. A method of mediating male fertility in a plant is
 XX provided that involves introducing an expression vector comprising a
 XX promoter operably linked to (I) into a plant where the exogenous gene
 XX impacts male fertility of the plant and (II) controls expression of the
 XX exogenous gene. A method of producing hybrid seeds is also provided. The

CC present sequence represents a nucleic acid sequence encoding an Ms45 male
 CC -cisise preferred regulatory region from Z. mays
 XX
 SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
 Query Match 100.0%; Score 1394; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAGGATCC 60
 Db 1 CCATGGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTCTTAAGGATCC 60

QY 61 CTCTCTCTGCTTATTAATGATCTGATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
 Db 61 CTCTCTCTGCTTATTAATGATCTGATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120

QY 121 CTCGATGTTTCACCTTCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCAT 180
 Db 121 CTCGATGTTTCACCTTCTCCACCTCGGTTGACATTTCTTGGATGCGGTGGTCCAT 180

QY 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCTTTTCGATGGCCCA 240
 Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGGCTTTTCGATGGCCCA 240

QY 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATGCTACCTTAT 300
 Db 241 CGAGACGTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCCAATGCTACCTTAT 300

QY 301 TTATATCTCTTATGATATTAATTTTGGAAAAATAGCAACTTATCTTTGTGTA 360
 Db 301 TTATATCTCTTATGATATTAATTTTGGAAAAATAGCAACTTATCTTTGTGTA 360

QY 361 GGGCTCAGACATGATTTTTCGTTAGGGCCAGAAAATGCGAGACAGCCATGCTTAGTG 420
 Db 361 GGGCTCAGACATGATTTTTCGTTAGGGCCAGAAAATGCGAGACAGCCATGCTTAGTG 420

QY 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAACAAAGTAACTATATCACT 480
 Db 421 TCCACTATTTGGCACTACCCAGAACAAAGATTTAAACAAAGTAACTATATCACT 480

QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACAGACATCTCTTAAAAA 540
 Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACAGACATCTCTTAAAAA 540

QY 541 CAAGCATATTTCCAAAAGACAATATGTTACAGTTTCAAAACATCTAAGACGACAAA 600
 Db 541 CAAGCATATTTCCAAAAGACAATATGTTACAGTTTCAAAACATCTAAGACGACAAA 600

QY 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGTTTCTTTTCATTTCTGTTATTTTGT 660
 Db 601 TTATATCGAAAGGTAAAGCTATGACGTTCAAGTTTCTTTTCATTTCTGTTATTTTGT 660

QY 661 ATTTGTTTATATACATTTTCTCTTACATAGAGTGATTTTCTTCGATTTTAA 720
 Db 661 ATTTGTTTATATACATTTTCTCTTACATAGAGTGATTTTCTTCGATTTTAA 720

QY 721 ATGACTATAAAGTCAATTTTATATAAGAGCAGCATGCTAGATTTCTCGTTCAAAAATC 780
 Db 721 ATGACTATAAAGTCAATTTTATATAAGAGCAGCATGCTAGATTTCTCGTTCAAAAATC 780

QY 781 TTTCTGATTTTTTAAAGAGTAAAGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
 Db 781 TTTCTGATTTTTTAAAGAGTAAAGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840

QY 841 TTCAAAAAAATAGTTATTTCTGTTTAAATAAGAAACACTTAAAGAAATAGAGT 900
 Db 841 TTCAAAAAAATAGTTATTTCTGTTTAAATAAGAAACACTTAAAGAAATAGAGT 900

QY 901 TGCAGAGTACCCCTAGAAATGTTTCCCAATAAATCAATCACTGTATTAATTTTG 960
 Db 901 TGCAGAGTACCCCTAGAAATGTTTCCCAATAAATCAATCACTGTATTAATTTTG 960

QY 961 GCCAGGCCATTAATTTATTTAAACCGAAATCGAAATCGGCGAAACCAATCTGAGCTAT 1020
 Db 961 GCCAGGCCATTAATTTATTTAAACCGAAATCGAAATCGGCGAAACCAATCTGAGCTAT 1020

QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAAGAAATCAGTTTAAAGTATGTTGCC 1080
 Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAAGAAATCAGTTTAAAGTATGTTGCC 1080

QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGTCTATAGGCTTACGTCA 1140
 Db 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGTCTATAGGCTTACGTCA 1140

QY 1141 GGTTCGAGCTCTGCTGTCATCTGACATGGAATCTACATGCTTGTAAACGTTGCTC 1200
 Db 1141 GGTTCGAGCTCTGCTGTCATCTGACATGGAATCTACATGCTTGTGTAAACGTTGCTC 1200

QY 1201 TTGTTTCATCTGTCGAAGCCTTGCTTATTTGAACCAAGAGATACCTATCCCAACAT 1260
 Db 1201 TTGTTTCATCTGTCGAAGCCTTGCTTATTTGAACCAAGAGATACCTATCCCAACAT 1260

QY 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAACCAATCAT 1320
 Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGACATATGTTTCTGAAACCAATCAT 1320

QY 1321 AAAAGTCAACAACGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380
 Db 1321 AAAAGTCAACAACGCTAGGTTCTCCGCTACTTCCCTCTCTCTGCGCATCTTTT 1380

QY 1381 CGTCCACGACATG 1394
 Db 1381 CGTCCACGACATG 1394

RESULT 3
 AAX07409
 ID AAX07409 standard; DNA; 1394 BP.

AC AAX07409;
 AC 08-JUN-1999 (first entry)
 DT
 XX
 DE Zea mays Ms45 male tissue-preferred regulatory region.
 XX
 KW Ms45; male; tissue-preferred; regulatory region; plant cells;
 KW plant tissue; differentiated; hybrid seed; fertility; ss.
 OS
 XX Zea mays.
 PN
 XX WO9859061-A1.
 XX
 PD 30-DEC-1998.
 XX
 PF 19-JUN-1998; 98WO-US012895.
 XX
 PR 23-JUN-1997; 97US-00880499.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PI
 XX Albreten MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
 PT WPI; 1999-105628/09.
 DR
 XX
 XX
 XX
 PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region.
 XX
 XX - useful in mediating plant fertility, especially hybrid seed production.
 PS
 XX Claim 3; Page 23-24; 39pp; English.
 CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue-preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocol/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably

Query Match 94.0%; Score 1311; DB 5; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTGCTCTATGAAAAAGATGAGTACATGTCCTATATCCGTTTCTTAAGGTC 60
 DB 1 CCATGCTGCTCTATGAAAAAGATGAGTACATGTCCTATATCCGTTTCTTAAGGTC 60

QY 61 CTTCCTTCTGCTTATTAAGTCTGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120
 DB 61 CTTCCTTCTGCTTATTAAGTCTGAATCGGGGTTTACAAAACTTCCAGGGTGCATGAT 120

QY 121 CTTCATGTTCCACTTCTCCACCTCGGCTGCACTTCTTGATGTCGGTGGTCCAT 180
 DB 121 CTTCATGTTCCACTTCTCCACCTCGGCTGCACTTCTTGATGTCGGTGGTCCAT 180

QY 181 CTGACCGAGGCCCATGAGACACCTTTGGGACACCCATCAAGGGCCTTTGGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATGAGACACCTTTGGGACACCCATCAAGGGCCTTTGGATGGCCCA 240

QY 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300
 DB 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCTCCACATGTCACCTATA 300

QY 301 TTATTAATCTTTAGATTTATTAATTTTGAATAAATAAATACTTATCTTTGTGTA 360
 DB 301 TTATTAATCTTTAGATTTATTAATTTTGAATAAATAAATACTTATCTTTGTGTA 360

QY 361 GGGCCTCAGCATGATTTTGGCTTAAAGGCCAGAAATGCGAGGACGAGCCATCTAGTG 420
 DB 361 GGGCCTCAGCATGATTTTGGCTTAAAGGCCAGAAATGCGAGGACGAGCCATCTAGTG 420

QY 421 TCCACTATTTGGCACTACCCAGAAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480
 DB 421 TCCACTATTTGGCACTACCCAGAAACAGATTTAAAAAATAACCAAGTAACTATCCACT 480

QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATCATTTAAACACAGATCCTTTAAAAAA 540
 DB 481 CGAAGCTATCATGTATGTTTAAAGAAACATCATTTAAACACAGATCCTTTAAAAAA 540

QY 541 CAAGCATATTTCCAAAGAGACAATATATGTTACAGTTTACAAACATCTAAGACGACAA 600
 DB 541 CAAGCATATTTCCAAAGAGACAATATATGTTACAGTTTACAAACATCTAAGACGACAA 600

QY 601 TTAATATCGAAAGGTAAGCTATGACGTTTCTTTTCACTTTGTGTAATTTGTT 660
 DB 601 TTAATATCGAAAGGTAAGCTATGACGTTTCTTTTCACTTTGTGTAATTTGTT 660

QY 661 ATGTTTAT 720
 DB 661 ATGTTTAT 720

QY 721 ATGACTATAAAGTCAATTTTATATTAAGAGCAGCATGTCGTAGATTCTCGTTCAAAAAT 780
 DB 721 ATGACTATAAAGTCAATTTTATATTAAGAGCAGCATGTCGTATATCTCGTTCAAAAAT 780

QY 781 TTTTGTATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTTGTATTTTAAAGCTAGTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840

QY 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAAGAAACACTTGAAGAAATAGAGT 900
 DB 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAAGAAACACTTGAAGAAATAGAGT 900

QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTACATCACTGTATATATATTTTG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTACATCACTGTATATATTTTG 960

QY 961 GCCAGCCCCCTAATATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGCCCCCTAATATTTTAAACCGAAATCGAGCGAAACCAATCTGAGCTAT 1020

QY 1021 TTCTTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCACTTTTAAGTATGTGCC 1080

DB 1021 TTCTTAGATTAGTAAAAAGGAGAGAGAGAAAGAAATCACTTTTAAGTATGTGCC 1080

QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGGTGCTACGTCA 1140
 DB 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGGTGCTACGTCA 1140

QY 1141 GGTTCGGCAGCTCTCGTGTATCTTCATGCGCATATCATATGTTTCAACCGTTTCGTC 1200
 DB 1141 GGTTCGGCAGCTCTCGTGTATCTTCATGCGCATATCATATGTTTCAACCGTTTCGTC 1200

QY 1201 TTGTTCCATGTCGCAAGCCCTTGCTTCTGAAACCAAGAGGATACCTACCTCCAAACAT 1260
 DB 1201 TTGTTCCATGTCGCAAGCCCTTGCTTCTGAAACCAAGAGATACCTACCTCCAAACAT 1260

QY 1261 CCATCTTACTCATGCAACTTTCATGCAACAGCAACGATATGTTTCTGTAAC 1311
 DB 1261 CCATCTTACTCATGCAACTTTCATGCAACAGCAACGATATGTTTCTGTAAC 1311

RESULT 5
 AAH76340
 ID AAH76340 standard; DNA; 255 BP.

AC AAH76340;
 XX
 XX 29-OCT-2001 (first entry)
 DT
 XX
 XX Z. mays Ms45 promoter fragment.
 DE
 XX Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; promoter; db.
 XX
 XX Zea mays.
 OS
 XX
 XX MO200160997-A2.
 PN
 XX
 XX 23-AUG-2001.
 PD
 XX
 XX 13-FEB-2001; 2001WO-US004527.
 PF
 XX
 XX 15-FEB-2000; 2000US-00504487.
 PR
 XX
 XX (PION-) PIONEER HI-BRED INT INC.
 PA
 XX
 XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
 PI
 XX
 XX WPI; 2001-514772/56.
 DR
 XX
 XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.
 PT
 XX
 XX Example 5; Fig 8; 50pp; English.
 PS
 XX
 XX The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (i) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (ii) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 CC
 XX
 XX Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

QY 1155 CGTGTATCTCATGCAATGCACTATCATAGCTTGTCAACGTTGCTTGTTCATGCTCC 1214
 DB 15 CGTGTATCTCATGCAATGCACTATCATAGCTTGTTCACCGTTGCTTGTTCATGCTCC 74

QY 1215 AAGCTTGGCTATTCTGAACCAAGAGATCTACTCCGAACATCCATTACTGAG 1274
 DB 75 AAGCTTGGCTATTCTGAACCAAGAGATCTACTCCGAACATCCATTACTGAG 134
 QY 1275 CAATTCATGCAACAACGACATATGTTCTCTGAAC 1311
 DB 135 CAATTCATGCAACAACGACATATGTTCTCTGAAC 171

RESULT 6

AAH76334
 ID AAH76334 standard; DNA; 158 BP.

XX AAH76334;

XX 29-OCT-2001 (first entry).

XX Z. mays Ms45 male tissue-preferred regulatory region fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; ds.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences

XX essential for initiating transcription of the Ms45 gene useful for

XX mediating fertility in a male plant.

XX Claim 5; Page 47; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (I)

XX comprising nucleotide sequences essential for initiating transcription of

XX the Ms45 gene. A method of mediating male fertility in a plant is

XX provided that involves introducing an expression vector comprising a

XX promoter operably linked to (I) into a plant where the exogenous gene

XX impacts male fertility of the plant and (I) controls expression of the

XX exogenous gene. A method of producing hybrid seeds is also provided. The

XX present sequence represents a DNA fragment -38 to -195 bases upstream of

XX the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region

XX nucleotide sequence

XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match

Best Local Similarity 8.0%; Score 111; DB 5; Length 158;

Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1201 TTGTTCCATGATCAAGCCTTGGCTATTCTGAACCAAGATACCTACTCCAAACAT 1260
 DB 48 TTGTTCCATGATCAAGCCTTGGCTATTCTGAACCAAGATACCTACTCCAAACAT 107
 QY 1261 CCATCTTACTGACCAACTTCCATGCAACGACATATGTTCTCTGAAC 1311
 DB 108 CCATCTTACTGACCAACTTCCATGCAACGACATATGTTCTCTGAAC 158

RESULT 7

AAH76336

ID AAH76336 standard; DNA; 50 BP.

XX AAH76336;

XX 29-OCT-2001 (first entry)

XX Z. mays Ms45 male tissue-preferred regulatory region fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

XX hybrid seed; ds.

XX Zea mays.

XX WO200160997-A2.

XX 23-AUG-2001.

XX 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences

XX essential for initiating transcription of the Ms45 gene useful for

XX mediating fertility in a male plant.

XX Claim 5; Page 47; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (I)

XX comprising nucleotide sequences essential for initiating transcription of

XX the Ms45 gene. A method of mediating male fertility in a plant is

XX provided that involves introducing an expression vector comprising a

XX promoter operably linked to (I) into a plant where the exogenous gene

XX impacts male fertility of the plant and (I) controls expression of the

XX exogenous gene. A method of producing hybrid seeds is also provided. The

XX present sequence represents a DNA fragment -72 to -111 bases upstream of

XX the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region

XX nucleotide sequence

XX Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;

XX Query Match 3.6%; Score 50; DB 5; Length 50;

XX Best Local Similarity 100.0%; Pred. No. 2.6e-13;

XX Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1229 CTGAACCAAGAGATACCTACTCCAAACATCCATCTTACTGATGCAAC 1278
 DB 1 CTGAACCAAGAGATACCTACTCCAAACATCCATCTTACTGATGCAAC 50

RESULT 8

AAH76337
 ID AAH76337 standard; DNA; 40 BP.
 XX AAH76337;
 XX 29-OCT-2001 (first entry)
 XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
 XX Ms45; male tissue; regulatory region; transcription; male fertility;
 XX hybrid seed; ds.
 XX Zea mays.
 XX WO200160997-A2.
 XX 23-AUG-2001.


```
XX 13-FEB-2001; 2001WO-US004527.
PF 15-FEB-2000; 2000US-00504487.
PR (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX MPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX PT essential for initiating transcription of the MS45 gene useful for
XX PT mediating fertility in a male plant.
XX
XX Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
XX CC comprising nucleotide sequences essential for initiating transcription of
XX CC the MS45 gene. A method of mediating male fertility in a plant is
XX CC provided that involves introducing an expression vector comprising a
XX CC promoter operably linked to (I) into a plant where the exogenous gene
XX CC impacts male fertility of the plant and (I) controls expression of the
XX CC exogenous gene. A method of producing hybrid seeds is also provided. The
XX CC present sequence represents a DNA fragment upstream of the TATA box of a
XX CC Z. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
XX Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
XX
XX Query Match 2.9%; Score 40; DB 5; Length 40;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-08;
XX Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1239 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 1278
XX 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAC 40
XX DB
XX
XX RESULT 9
XX AAH76335
XX ID AAH76335 standard; DNA; 30 BP.
XX
XX AC AAH76335;
XX
XX DT 29-OCT-2001 (first entry)
XX
XX DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX KM hybrid seed; db.
XX
XX OS Zea mays.
XX
XX PN WO200160997-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 13-FEB-2001; 2001WO-US004527.
XX
XX PR 15-FEB-2000; 2000US-00504487.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX MPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
XX PT essential for initiating transcription of the MS45 gene useful for
XX PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
```

```
CC The invention provides a male tissue-preferred regulatory region (I)
CC CC comprising nucleotide sequences essential for initiating transcription of
CC CC the MS45 gene. A method of mediating male fertility in a plant is
CC CC provided that involves introducing an expression vector comprising a
CC CC promoter operably linked to (I) into a plant where the exogenous gene
CC CC impacts male fertility of the plant and (I) controls expression of the
CC CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC CC present sequence represents a DNA fragment -152 to -181 bases upstream of
CC CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
XX CC nucleotide sequence
XX
XX SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;
XX
XX Query Match 2.2%; Score 30; DB 5; Length 30;
XX Best Local Similarity 100.0%; Pred. No. 0.0011;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1179 CATGCTGTTCACCGTTCGCTGTTCGA 1208
XX 1 CATGCTGTTCACCGTTCGCTGTTCGA 30
XX DB
XX
XX RESULT 10
XX ABL40437/C
XX ID ABL40437 standard; DNA; 849 BP.
XX
XX AC ABL40437;
XX
XX DT 10-JUN-2002 (first entry)
XX
XX DE Sunflower chitinase promoter sequence.
XX
XX KM Sunflower; chitinase; lipid transfer protein; LTP; transcription;
XX KM sclerotinia; genetic engineering; disease resistance; insecticide;
XX KM antifungal; viricide; nematocide; antimicrobial; antibacterial;
XX KM gene therapy; gene; promoter; ds.
XX
XX OS Helianthus annuus.
XX
XX FH Key Location/Qualifiers
XX FT CAAAT_signal 723..726
XX FT TATA_signal 807..811
XX FT /*tag= a
XX FT /*tag= b
XX
XX PN WO200214502-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 08-AUG-2001; 2001WO-US041629.
XX
XX PR 11-AUG-2000; 2000US-0224603P.
XX
XX PR 07-AUG-2001; 2001US-00923844.
XX
XX PA (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Bao Z, Lu G;
XX MPI; 2002-269194/31.
XX
XX New genes and promoters from Helianthus annuus, for genetically
XX PT manipulating plants to enhance disease resistance or resistance to
XX PT pathogen, and for producing proteins for controlling plant diseases
XX PT caused by these pathogens.
XX
XX Claim 1; Fig 4; 89pp; English.
XX
XX The invention relates to sclerotinia-inducible genes and promoters capable
XX CC of initiating transcription from sunflower. The polynucleotides are
XX CC useful for genetically manipulating plants to enhance disease resistance,
XX CC or resistance to insect, fungal pathogens, viruses or nematodes. The
XX CC nucleotide sequences are also useful as genetic markers in disease-
XX CC resistance breeding programs. The polynucleotides are also useful for
```


CC producing proteins useful in compositions. The compositions are useful
CC reducing or eliminating damage to plants caused by plant pathogens. The
CC compositions find use in agricultural and pharmaceutical compositions as
CC antifungal and antimicrobial agents. For agricultural purposes, the
CC compositions may be used in sprays for control of plant disease. As
CC pharmaceutical compositions, the agents are useful as antibacterial and
CC antimicrobial treatments. The present sequence represents the promoter
CC fragment of the sunflower chitinase gene
XX
SQ Sequence 849 BP; 295 A; 139 C; 135 G; 280 T; 0 U; 0 Other;

Query Match 1.6%; Score 23; DB 6; Length 849;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 854 AGTTATTTCTCTTATAAAT 876
DB 556 AGTTATTTCTCTTATAAAT 534

RESULT 11
ABV56505
ID ABV56505 standard; cDNA; 333 BP.
XX
AC ABV56505;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 56496.
XX
KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183119P.
XX
PR 16-MAR-2000; 2000US-0189662P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JR;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 10900; 11750pp; English.
XX
SQ The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 333 BP; 75 A; 69 C; 46 G; 143 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 5; Length 333;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 TTACAAAAAATTCACGGGT 113
DB 89 TTACAAAAAATTCACGGGT 109

RESULT 12
ABA58867/c
ID ABA58867 standard; DNA; 459 BP.
XX
AC ABA58867;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #7172.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-0060840P.
XX
PR 03-AUG-2000; 2000US-0063236P.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
PS Claim 1; SEQ ID NO 7172; 639pp + Sequence Listing; English.
XX
SQ The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98

RESULT 13
AAI38581/c
ID AAI38581 standard; DNA; 459 BP.


```
XX AC AA138581;
XX DT 17-OCT-2001 (first entry)
XX DE Probe #7267 used to measure gene expression in human placenta sample.
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200157272-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000663.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-48897/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX PS Claim 25; SEQ ID NO 7267; 654bp; English.
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX SO Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match      1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
DB      118 AAAGGAGAGAGAGAGAGA 98

RESULT 14
ID AAK32768/c
AC AAK32768 standard; DNA; 459 BP.
XX AC AAK32768;
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 7325.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX
```

```
PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488900/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 7325; 658bp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention
XX SO Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match      1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1037 AAAGGAGAGAGAGAGAGA 1057
DB      118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
ID AAK07027/c
AC AAK07027 standard; DNA; 459 BP.
XX AC AAK07027;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 7018.
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX KW ss.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
```


DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
XX
PS Example 4; SEQ ID NO 7018; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 459 BP; 86 A; 149 C; 70 G; 154 T; 0 U; 0 Other;

Query Match 1.5%; Score 21; DB 4; Length 459;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
|||
Db 118 AAAGGAGAGAGAGAGAGA 98

Search completed: September 16, 2005, 13:08:39
Job time : 584.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 19:42:02 ; Search time 176.5 Seconds
(without alignments)
12923.341 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 cccttgctgcctcctatgaaa.....cttttcgccaccaccatg 1394

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued, Patente NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB. seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB. seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB. seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB. seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS COMB. seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	3	US-08-880-499-1
2	1311	94.0	1394	4	US-08-880-499-2
3	21	1.5	50263	4	US-09-949-016-13563
4	20	1.4	601	4	US-09-949-016-124481
5	20	1.4	601	4	US-09-949-016-147075
6	20	1.4	2523	2	US-08-410-784A-3
7	20	1.4	21679	4	US-09-949-016-15250
8	20	1.4	65424	4	US-09-949-016-12426
9	20	1.4	374159	4	US-09-949-016-15868
10	19	1.4	601	4	US-09-949-016-17933
11	19	1.4	601	4	US-09-949-016-17934
12	19	1.4	601	4	US-09-949-016-41823
13	19	1.4	601	4	US-09-949-016-41824
14	19	1.4	601	4	US-09-949-016-145344
15	19	1.4	601	4	US-09-949-016-145345
16	19	1.4	766	4	US-09-328-352-2409
17	19	1.4	1584	1	US-08-247-808A-10
18	19	1.4	1584	1	US-08-453-942-10
19	19	1.4	1584	2	US-08-926-885A-10
20	19	1.4	1584	5	PCT-US94-05290-10
21	19	1.4	1833	1	US-08-365-470-2
22	19	1.4	3834	3	US-09-209-668-18
23	19	1.4	3834	4	US-09-949-016-71
24	19	1.4	3834	1	US-08-365-470-1
25	19	1.4	3858	2	US-08-344-155C-98
26	19	1.4	3858	3	US-09-009-490A-88
27	19	1.4	3862	4	US-09-949-016-4401

C 28	19	1.4	3863	3	US-08-482-073-1	Sequence 1, Appl1
C 29	19	1.4	3863	6	5217870-1	Patent No. 5217870
C 30	19	1.4	3863	6	5217870-1	Patent No. 5217870
C 31	19	1.4	9370	1	US-08-320-559-27	Sequence 27, Appl1
C 32	19	1.4	9370	3	US-08-545-860D-27	Sequence 27, Appl1
C 33	19	1.4	9370	5	PCT-US94-04496-27	Sequence 27, Appl1
C 34	19	1.4	9391	1	US-08-320-559-25	Sequence 25, Appl1
C 35	19	1.4	9391	3	US-08-545-860D-25	Sequence 25, Appl1
C 36	19	1.4	9391	5	PCT-US94-04496-25	Sequence 25, Appl1
C 37	19	1.4	15384	4	US-09-949-016-11813	Sequence 11813, A
C 38	19	1.4	15413	4	US-09-949-016-16143	Sequence 16143, A
C 39	19	1.4	71645	4	US-09-949-016-12126	Sequence 12126, A
C 40	19	1.4	71651	4	US-09-949-016-17258	Sequence 17258, A
C 41	19	1.4	105189	4	US-09-949-016-13029	Sequence 13029, A
C 42	19	1.4	133559	4	US-09-949-016-15845	Sequence 15845, A
C 43	19	1.4	144158	4	US-09-949-016-11755	Sequence 11755, A
C 44	19	1.4	144158	4	US-09-949-016-12936	Sequence 12936, A
C 45	19	1.4	228896	4	US-09-949-016-17127	Sequence 17127, A

ALIGNMENTS

RESULT 1
US-08-880-499-1

; Sequence 1, Application US/08880499
; Patent No. 6037523

; GENERAL INFORMATION:

; APPLICANT: Albertson, Marc C.

; APPLICANT: Fox, Tim W.

; APPLICANT: Carl, Gary W.

; APPLICANT: Huffman, Gary A.

; APPLICANT: Kendall, Timmy L.

; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.

; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

; CITY: Johnston

; STATE: Iowa

; COUNTRY: USA

; ZIP: 50131

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/880,499

; FILING DATE: CONCURRENTLY HERewith

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Sweeney, Patricia A.

; REGISTRATION NUMBER: 32,733

; REFERENCE/DOCKET NUMBER: 0578

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (515) 248-4800

; TELEFAX: (515) 248-4844

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1394 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-880-499-1

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGCTCTCTATGAAAAAGATGATCATATGTCATATCCGTTTCTTAGGCTC 60
Db 1 CCATGCTCTCTATGAAAAAGATGATCATATGTCATATCCGTTTCTTAGGCTC 60
QY 61 CTTCTTCTGCTTATTACTGATCGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
Db 61 CTTCTTCTGCTTATTACTGATCGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
QY 121 CTCATGTTTCACCTTCCGACCTCGGTGACATTTCTTGGATGTCGGTGGTCCCAT 180
Db 121 CTCATGTTTCACCTTCCGACCTCGGTGACATTTCTTGGATGTCGGTGGTCCCAT 180
QY 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCCATCAAGGCGCTTTCGATGGCCCA 240
Db 181 CTGACCGAGGCCCATGACACCTTTCGGGACACCCATCAAGGCGCTTTCGATGGCCCA 240
QY 241 CGAGAGCTATCGGGTCTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTCTGTGATCCAGGGGATATATGTCCTCCCAACATGTCACCTATA 300
QY 301 TTATTTATCTTTAGATTATTTAATTTTGGAAAAATAACAACCTTATCTTTGGTA 360
Db 301 TTATTTATCTTTAGATTATTTAATTTTGGAAAAATAACAACCTTATCTTTGGTA 360
QY 361 GGGCTTCAGCATGATTTTCCTTAGGGCCAGAAAAATGCGAGGACAGCCATGCTAGTG 420
Db 361 GGGCTTCAGCATGATTTTCCTTAGGGCCAGAAAAATGCGAGGACAGCCATGCTAGTG 420
QY 421 TCCACTTATGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCACT 480
Db 421 TCCACTTATGGCACTACCCAGAACAGATTTAAAAAATACCAAGTAACTATATCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCTCTTAAAAA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCATTTAAACACGATCCTCTTAAAAA 540
QY 541 CAAGCATATTCGAAAGAGACAAATATGTTACAGTTTACCAACATCTAGAGCGACAA 600
Db 541 CAAGCATATTCGAAAGAGACAAATATGTTACAGTTTACCAACATCTAGAGCGACAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGTTTTCTTTTCATTTCTGTATTTTGT 660
Db 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGTTTTCTTTTCATTTCTGTATTTTGT 660
QY 661 ATGTTTTTATATACATTTCTTCTCTTATACATAGATGATTTTCTCCGATTTTAAA 720
Db 661 ATGTTTTTATATACATTTCTTCTCTTATACATAGATGATTTTCTCCGATTTTAAA 720
QY 721 ATGACTATAAAGTCAATTTTATATAAGAGACGCAATGCTAGATTCCTGTTCAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATAAGAGACGCAATGCTAGATTCCTGTTCAAAAATC 780
QY 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTTGATTTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTTGATTTT 840
QY 841 TTTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACCTTAGAAAAATAGAGT 900
Db 841 TTTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACCTTAGAAAAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACATCACTGTATATTAATTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACATCACTGTATATTAATTG 960
QY 961 GCGAGGCCCATTAATTTTAAACCGAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
Db 961 GCGAGGCCCATTAATTTTAAACCGAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGCTATGCTCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGCTATGCTCC 1080
QY 1081 TGAGATGTGCGGTTTGGCAACGATAGCCAGTAAATCATAGTATGCTAGTCCCTAGCTCA 1140

Db 1081 TGAGATGTGCGGTTTGGCAACGATAGCCAGCTAAATCAAGCTATAGTGTCTAGCTCA 1140
QY 1141 GGTTCGGACGCTCTGTGTCATCTCATGAGCATACATGCTGTTTCAACCGTTGCTC 1200
Db 1141 GGTTCGGACGCTCTGTGTCATCTCATGAGCATACATGCTGTTTCAACCGTTGCTC 1200
QY 1201 TTGTTTCATGTCGCAAGCGCTTGCTATTGTAACCAAGAGGATPACTACTCCCAACAT 1260
Db 1201 TTGTTTCATGTCGCAAGCGCTTGCTATTGTAACCAAGAGGATPACTACTCCCAACAT 1260
QY 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACCAATCCATT 1320
Db 1261 CCATCTTACTCATGCACTTCCATGCAACACGACATATGTTTCTGAAACCAATCCATT 1320
QY 1321 AAAAGTCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAAGTCAACAAGCTAGGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
QY 1381 CGTCCACGACGATG 1394
Db 1381 CGTCCACGACGATG 1394

RESULT 2
US-08-880-499-2
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Albertson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnat W.
; APPLICANT: Huffman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; NUMBER OF INVENTIONS: 2
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; CITY: Box 1000
; STREET: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880 499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-2
Query Match 94.0%; Score 1311; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATGCTCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGGCTCC 60
DB 1 CCAATGGGTCTCTATGAAAAAGATGATACATGCTCTATATCCGTTTCTTAGGCTCC 60
QY 61 CTTCCTTCGCTTATTAAGTACTGATCGGGGTTACAAAAAAGTTCACCGGGTGCATGAT 120
DB 61 CTTCCTTCGCTTATTAAGTACTGATCGGGGTTACAAAAAAGTTCACCGGGTGCATGAT 120
QY 121 CTTCATGTTCCACTTCTCCCACTCGGGTGCACATTTCTGGATGTCGGTGGTCCAT 180
DB 121 CTTCATGTTCCACTTCTCCCACTCGGGTGCACATTTCTGGATGTCGGTGGTCCAT 180
QY 181 CTGACCGAGGCCCATGAGACACTTTCGGGACACCCATCAAGGGCTTTGGATGCCCCA 240
DB 181 CTGACCGAGGCCCATGAGACACTTTCGGGACACCCATCAAGGGCTTTGGATGCCCCA 240
QY 241 CGAGACGTATCGGGTCTGCTGATCCAGGGGATATATGTCGCCCACAATGCTACCTATA 300
DB 241 CGAGACGTATCGGGTCTGCTGATCCAGGGGATATATGTCGCCCACAATGCTACCTATA 300
QY 301 TTATTAATCTTATGATATTAATTAATTTTGAATAAACAATTAATCTTTTGTGTA 360
DB 301 TTATTAATCTTATGATATTAATTAATTTTGAATAAACAATTAATCTTTTGTGTA 360
QY 361 GGGGCTCAGATAGATTTTGGCTTAGGGCCCAAGAAATGCGAGACCAAGCATGCTTAGTG 420
DB 361 GGGGCTCAGATAGATTTTGGCTTAGGGCCCAAGAAATGCGAGACCAAGCATGCTTAGTG 420
QY 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAATAACAAAGTAACTAATCACT 480
DB 421 TCCACTATTGGCACTACCCAGAACAGATTTAAATAAATAACAAAGTAACTAATCACT 480
QY 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAATAA 540
DB 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACACGATCCTTTAAATAA 540
QY 541 CAAAGCTATTTGCAAAAGACAAATTAATGTTACAGTTTAAACATCTAAGACGCAAA 600
DB 541 CAAAGCTATTTGCAAAAGACAAATTAATGTTACAGTTTAAACATCTAAGACGCAAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTTCAATTTTCTTTTCAATCTTGTATTTGTT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTTCAATTTTCTTTTCAATCTTGTATTTGTT 660
QY 661 ATTGTTTTATATACATTTTCTTCTTACATAGAGATTTTCTCCGATTTTAA 720
DB 661 ATTGTTTTATATACATTTTCTTCTTACATAGAGATTTTCTCCGATTTTAA 720
QY 721 ATGACTATTAAGCTATTTTATATAGAGACGCAATGCTGATGTTCTGTTCAAAAATC 780
DB 721 ATGACTATTAAGCTATTTTATATAGAGACGCAATGCTGATGTTCTGTTCAAAAATC 780
QY 781 TTTCGATTTTTTAAAGAGTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
DB 781 TTTCGATTTTTTAAAGAGTAGTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
QY 841 TTCAAAAAAATAGTTTATTTCTTTTATTAATAAGAAACACTTAGAAAAATAGAGT 900
DB 841 TTCAAAAAAATAGTTTATTTCTTTTATTAATAAGAAACACTTAGAAAAATAGAGT 900
QY 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATAAATTAATCACTGCTGTAATTTTGG 960
DB 901 TGGCAACTAGCCCTAGAAATGTTTCCCAATAAATTAATCACTGCTGTAATTTTGG 960
QY 961 GCGAGCCCATTAATTAATTAACCGAAATCGAAGTCAAGCAATCTGAGCTAT 1020
DB 961 GCGAGCCCATTAATTAATTAACCGAAATCGAAGTCAAGCAATCTGAGCTAT 1020
QY 1021 TTCTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCTGCC 1080
DB 1021 TTCTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCTGCC 1080

QY 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGTGTCTAGTCA 1140
DB 1081 TGAGATGTCGGTTTGGCAAGATAGCCACCGTAATCATAGCTCATAGTGTCTAGTCA 1140
QY 1141 GGTTCGGAGCTCTCGTGTATCTCATGATGAGATACATGCTTTGTAACCGTTGTC 1200
DB 1141 GGTTCGGAGCTCTCGTGTATCTCATGATGAGATACATGCTTTGTAACCGTTGTC 1200
QY 1201 TTGTTCCATGTCGAAGCCCTGCTATTTGTAACCAAGAGATACCTACTCCAAACAT 1260
DB 1201 TTGTTCCATGTCGAAGCCCTGCTATTTGTAACCAAGAGATACCTACTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTTCATGCAACAGCAATATGTTCTCTGAAC 1311
DB 1261 CCATCTTACTCATGCAACTTTCATGCAACAGCAATATGTTCTCTGAAC 1311

RESULT 3
US-09-949-016-13563
; Sequence 13563, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13563
; LENGTH: 50263
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13563

Query Match 1.5%; Score 21; DB 4; Length 50263;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 861 TTTCCTTTATTAATAAGAA 881
DB 40441 TTTCCTTTATTAATAAGAA 40461

RESULT 4
US-09-949-016-124481/c
; Sequence 124481, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124481
; LENGTH: 601
; TYPE: DNA

ORGANISM: Human
US-09-949-016-124481

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCTTCTT 649
|||||
DB 363 GATTTTCTTTTCTTCTT 344

RESULT 5

US-09-949-016-147075/c
; Sequence 147075, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 147075
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-147075

Query Match 1.4%; Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 817 TTTCCTTCAAGATTTTGA 836
|||||
DB 353 TTTCCTTCAAGATTTTGA 334

RESULT 6

US-08-410-784A-3/c
; Sequence 3, Application US/08410784A
; Patent No. 5912413
; GENERAL INFORMATION:
; APPLICANT: MYERS, ALAN M.
; APPLICANT: JAMES, MARTHA G.
; TITLE OF INVENTION: ISOLATION OF SUV, A STARCH DEBRANCHING
; TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/410,784A
FILING DATE: 24-MAR-1995
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Ph.D., Holliday C
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: ISU-002XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-2290
TELEFAX: 617-451-0313
TELEX:

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
US-08-410-784A-3

Query Match 1.4%; Score 20; DB 2; Length 2523;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 849 AAATTAGTTATTTCCTT 868
|||||
DB 206 AAATTAGTTATTTCCTT 187

RESULT 7

US-09-949-016-15250
; Sequence 15250, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15250
; LENGTH: 21679
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15250

Query Match 1.4%; Score 20; DB 4; Length 21679;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 630 GATTTTCTTTTCTTCTT 649
|||||
DB 4880 GATTTTCTTTTCTTCTT 4899

RESULT 8

US-09-949-016-12426/c
; Sequence 12426, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.


```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12426
; LENGTH: 65424
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(65424)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12426

Query Match          1.4%: Score 20; DB 4; Length 65424;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1041 GGAGAGAGAGAGAGAGAAAT 1060
Db      50936 GGAGAGAGAGAGAGAGAAAT 50917

RESULT 9
US-09-949-016-15868
; Sequence 15868, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15868
; LENGTH: 374159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15868

Query Match          1.4%: Score 20; DB 4; Length 374159;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      817 TTTCCTTCAAGATTTTGA 836
Db      265055 TTTCCTTCAAGATTTTGA 265074

RESULT 10
US-09-949-016-17933/C
; Sequence 17933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17933
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17933

Query Match          1.4%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      449 TTTAAAAAATACCAAG 467
Db      361 TTTAAAAAATACCAAG 343

RESULT 11
US-09-949-016-17934/C
; Sequence 17934, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17934
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17934

Query Match          1.4%: Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      449 TTTAAAAAATACCAAG 467
Db      394 TTTAAAAAATACCAAG 376

RESULT 12
US-09-949-016-41823/C
; Sequence 41823, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
```



```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41823
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41823
```

```
Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      449 TTTAAAAAATTAACCAAG 467
      |||||
Db      361 TTTAAAAAATTAACCAAG 343
```

```
RESULT 13
US-09-949-016-41824/c
; Sequence 41824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41824
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-41824
```

```
Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      449 TTTAAAAAATTAACCAAG 467
      |||||
Db      394 TTTAAAAAATTAACCAAG 376
```

```
RESULT 14
US-09-949-016-145344/c
; Sequence 145344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145344
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145344
```

```
Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1038 AAGGAGAGAGAGAGGAG 1056
      |||||
Db      561 AAGGAGAGAGAGAGGAG 543
```

```
RESULT 15
US-09-949-016-145345/c
; Sequence 145345, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 145345
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145345
```

```
Query Match          1.4%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1038 AAGGAGAGAGAGAGGAG 1056
      |||||
Db      553 AAGGAGAGAGAGAGGAG 535
```

```
Search completed: September 16, 2005, 08:14:15
Job time : 179.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 03:01:45 ; Search time 1291.5 Seconds
(without alignments)
7191.347 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394

Sequence: 1 ccattgctcctcctatgaaaa.....cttttcgtccaccaccatcg 1394

Scoring table: Oligo NUC
Gapop 60.0 , Gapext 60.0

Searched: 7389322 seqs, 333128559 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10G_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10H_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US10I_PUBCOMB.seq:*
21: /cgn2_6/prodata/1/pubpna/US10J_PUBCOMB.seq:*
22: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/prodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1394	100.0	1394	20	US-10-713-381-1
2	1311	94.0	1394	20	US-10-713-381-2
3	157	11.3	255	20	US-10-713-381-9
4	111	8.0	158	20	US-10-713-381-3
5	70	5.0	1663	20	US-10-739-930-1857
6	50	3.6	50	20	US-10-713-381-5
7	40	2.9	40	20	US-10-713-381-6

	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
C	9	23	2	2	30	2	2	30	20	US-10-713-381-4	Sequence 4, Appl1																											
C	10	21	1.6	849	9	US-09-923-844B-5	Sequence 5, Appl1																															
C	11	22	1.6	194945	17	US-10-085-117-355	Sequence 355, App																															
C	12	21	1.5	333	20	US-10-357-930-56524	Sequence 56524, A																															
C	13	21	1.5	356	20	US-10-425-115-183040	Sequence 183040,																															
C	14	21	1.5	386	20	US-10-425-115-117183	Sequence 117183,																															
C	15	21	1.5	459	9	US-09-864-761-11465	Sequence 11465, A																															
C	16	21	1.5	1203	20	US-10-425-115-51470	Sequence 51470, A																															
C	17	21	1.5	1637	18	US-10-425-114-7365	Sequence 7365, Ap																															
C	18	21	1.5	1929	18	US-10-424-599-21199	Sequence 21199, A																															
C	19	21	1.5	9265	15	US-10-311-455-2125	Sequence 2125, Ap																															
C	20	21	1.5	9265	18	US-10-240-454-49	Sequence 49, Appl																															
C	21	21	1.5	18817	15	US-10-240-485-47	Sequence 47, Appl																															
C	22	20	1.4	255	20	US-09-984-428-240	Sequence 240, App																															
C	23	20	1.4	257	18	US-10-425-115-114007	Sequence 114007,																															
C	24	20	1.4	260	20	US-10-424-599-119819	Sequence 119819,																															
C	25	20	1.4	311	20	US-10-357-930-60745	Sequence 60745, A																															
C	26	20	1.4	605	13	US-10-425-115-176795	Sequence 176795,																															
C	27	20	1.4	605	13	US-10-027-632-139460	Sequence 139460,																															
C	28	20	1.4	605	17	US-10-027-632-139461	Sequence 139461,																															
C	29	20	1.4	605	17	US-10-027-632-139460	Sequence 139460,																															
C	30	20	1.4	680	20	US-10-027-632-139461	Sequence 139461,																															
C	31	20	1.4	680	20	US-10-363-345A-4725	Sequence 4725, Ap																															
C	32	20	1.4	680	21	US-10-363-345A-4725	Sequence 4725, Ap																															
C	33	20	1.4	680	21	US-10-363-483A-4725	Sequence 4726, Ap																															
C	34	20	1.4	680	21	US-10-363-483A-4726	Sequence 4726, Ap																															
C	35	20	1.4	821	20	US-10-425-115-34238	Sequence 34238, A																															
C	36	20	1.4	973	20	US-10-425-115-63769	Sequence 63769, A																															
C	37	20	1.4	1261	20	US-10-425-115-134230	Sequence 134230,																															
C	38	20	1.4	1518	20	US-10-425-115-35698	Sequence 35698, A																															
C	39	20	1.4	51198	13	US-10-087-192-1729	Sequence 1729, Ap																															
C	40	20	1.4	54060	21	US-10-741-600-17886	Sequence 17886, A																															
C	41	20	1.4	72019	13	US-10-741-600-17775	Sequence 17775, A																															
C	42	20	1.4	91871	13	US-10-087-192-235	Sequence 235, App																															
C	43	20	1.4	148497	20	US-10-723-860-1744	Sequence 1744, Ap																															
C	44	20	1.4	148497	22	US-10-755-149-675	Sequence 675, App																															
C	45	20	1.4	337344	19	US-10-388-838-58	Sequence 58, Appl																															
C	45	20	1.4	465237	9	US-09-933-267A-1	Sequence 1, Appl1																															

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/860,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1394; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGCTCTCTATGAAAAGATGATGATGCTCTATATCCGTTTCTTAGGCTCC 60

Db 1 CCAATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTTCTGCTTATTAAGTACTGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
Db 61 CTTCTTCTGCTTATTAAGTACTGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
Qy 121 CTCATGTTTCACTTCTCCACCTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCTCCACCTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTTGGATGGCCCA 240
Qy 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACCTATA 300
Db 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACCTATA 300
Qy 301 TTATATTTCTTATGATATTAATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATTTCTTATGATATTAATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGCCTCAGCATAGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
Db 361 GGGCCTCAGCATAGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGCATGTCTAGTG 420
Qy 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAATAACCAAGTATCTATCTACT 480
Db 421 TCCACTATTTGGCATACCCAGAAACAAGTTTAAATAACCAAGTATCTATCTACT 480
Qy 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCCCTTAAAAAA 540
Db 481 CGAAGCTATCATGTATGTTTAAAGAAACATCTATTAACCAAGATCCCTTAAAAAA 540
Qy 541 CAAGCATATTTGCAAGAGAGCAAAATATGTATACAGTTTACAAACATCTAAGACGACAA 600
Db 541 CAAGCATATTTGCAAGAGAGCAAAATATGTATACAGTTTACAAACATCTAAGACGACAA 600
Qy 601 TTATATCGAAGAGTATGCTATGACGTTCCAGATTTTCTTTTCAATCTTTGTATTTGTT 660
Db 601 TTATATCGAAGAGTATGCTATGACGTTCCAGATTTTCTTTTCAATCTTTGTATTTGTT 660
Qy 661 ATTGTTTTATATACATTTTCTCTCTTCAATAGAGTATTTCTTCCGATTTTATATA 720
Db 661 ATTGTTTTATATACATTTTCTCTCTTCAATAGAGTATTTCTTCCGATTTTATATA 720
Qy 721 ATGACTATATAAGTCAATTTTATATATAAGACGCAATGTGTGATTTCTGTTCAAAAAATC 780
Db 721 ATGACTATATAAGTCAATTTTATATAAGACGCAATGTGTGATTTCTGTTCAAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTAAAGCTAGTTTGGCAACCTGTCTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGAGT 900
Db 841 TTCAAAAAAAATAGTTTATTTCTCTTATTAATAAGAAAACTTGAATAATAGAGT 900
Qy 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTATATATTTG 960
Db 901 TGGCAAGCTAGCCCTAGAAATGTTTCCCAATAAATACATCACTGTATATATTTG 960
Qy 961 GCGAGGCCCATTAATTTTAAACGAAATCTGAATCGAGCGAAACCAAACTGAGACTAT 1020
Db 961 GCGAGGCCCATTAATTTTAAACGAAATCTGAATCGAGCGAAACCAAACTGAGACTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTGCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTCAATGTGCC 1080
Qy 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
Db 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140

Db 1081 TGAAGTGTGCGGTTTGGCAACGATAGCCACCGTAATCATAGCTCATAGGTGCTTACGTCA 1140
Qy 1141 GGTGGGCAAGCTCTGTGTCTATCTACATGGCATATACATGCTTTTAAACCGTTGCTC 1200
Db 1141 GGTGGGCAAGCTCTGTGTCTATCTACATGGCATATACATGCTTTTAAACCGTTGCTC 1200
Qy 1201 TTGTTTCATGCTCAAGCCCTTCTATTTGTAACCAAGAGGATACCTACTCCCAACAT 1260
Db 1201 TTGTTTCATGCTCAAGCCCTTCTATTTGTAACCAAGAGGATACCTACTCCCAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGTAACCAATTCAT 1320
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACACGCAATATGTTTCTGTAACCAATTCAT 1320
Qy 1321 AAAAGATCAACAGCTAGGTTCTCCGTTAGCTTCCCTCTCTCTGCGGATCTTTT 1380
Db 1321 AAAAGATCAACAGCTAGGTTCTCCGTTAGCTTCCCTCTCTCTGCGGATCTTTT 1380
Qy 1381 CGTCCACGACATG 1394
Db 1381 CGTCCACGACATG 1394

RESULT 2
US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 94.0%; Score 1311; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Db 1 CCATGCTCTCTATGAAAAAGATGAGTACAAATGTGTCTATATCCGTTTCTTAGGCTCC 60
Qy 61 CTTCTTGGCCTTATTAAGTACTGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
Db 61 CTTCTTGGCCTTATTAAGTACTGAATCGGGGTTTACAAAACTTCCACGGGTGATGAT 120
Qy 121 CTCATGTTTCACTTCTCCACCTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Db 121 CTCATGTTTCACTTCTCCACCTCGGGTGGACATTTCTTGGATGTGGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTTGGATGGCCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCTTTTGGATGGCCCA 240
Qy 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACCTATA 300
Db 241 CGAGAGGTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCTACCTATA 300
Qy 301 TTATATTTCTTATGATATTAATTTTGAAGAAATACAACTTATCTTTGTGTA 360
Db 301 TTATATTTCTTATGATATTAATTTTGAAGAAATACAACTTATCTTTGTGTA 360


```

Db      301 TTAATTATCTTTAGATATTTATTTATTTTGGAAAAATTAACAATTAATCTTTTGGTA 360
Qy      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAGTG 420
Db      361 GGGCCTCAGCATAGATTTTCGCTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAGTG 420
Qy      421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATTAACAAAGTAATCATCACT 480
Db      421 TCCACTATTGGCACTACCCAGAACAGATTTAAAAAATTAACAAAGTAATCATCACT 480
Qy      481 CGAAAGCTATCATGATATGTTTAAAGAAACATGATTTAAACCAAGTCCTTTAAAAA 540
Db      481 CGAAAGCTATCATGATATGTTTAAAGAAACATGATTTAAACCAAGTCCTTTAAAAA 540
Qy      541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAGAGCGACAA 600
Db      541 CAAGCATATTTGCAAAAGACAAATTAATGTTACAGTTTACAAACATCTAGAGCGACAA 600
Qy      601 TTTATTCGAAAGGTAAGCTATGACGTTGAGATTTTCTTTTCACTTCTGTTATTTGTT 660
Db      601 TTTATTCGAAAGGTAAGCTATGACGTTGAGATTTTCTTTTCACTTCTGTTATTTGTT 660
Qy      661 ATTGTTTATATACATTTTCTTCTTACATAGAGTATTTCTTCCGATTTTATAA 720
Db      661 ATTGTTTATATACATTTTCTTCTTACATAGAGTATTTCTTCCGATTTTATAA 720
Qy      721 ATGACTATTAAGTCATTTTATATTAAGAGCAGCATGCTGATAGTTCTGTTCAAAAATC 780
Db      721 ATGACTATTAAGTCATTTTATATTAAGAGCAGCATGCTGATAGTTCTGTTCAAAAATC 780
Qy      781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Db      781 TTTCTGATTTTAAAGAGTATGTTGGCAACCTGTTTCTTCAAAAGATTTGATTTT 840
Qy      841 TTCAAAAAAATAGTTTATTTCTCTTATAAATAGAAAAACATAGAAAAATAGAGT 900
Db      841 TTCAAAAAAATAGTTTATTTTCTTCTTATAAATAGAAAAACATAGAAAAATAGAGT 900
Qy      901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGATTAATTAATTTG 960
Db      901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACAATCACTGATTAATTAATTTG 960
Qy      961 GCCAGCCCCATTAATTTTAAACCGAAATCGAATCGAGGAAACCAAAATCTGAGCTAT 1020
Db      961 GCCAGCCCCATTAATTTTAAACCGAAATCGAATCGAGGAAACCAAAATCTGAGCTAT 1020
Qy      1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTGTC 1080
Db      1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTGTC 1080
Qy      1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTATCATAGTATAGTGTGCTTACGTCA 1140
Db      1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTATCATAGTATAGTGTGCTTACGTCA 1140
Qy      1141 GGTTCGGCAGCTCTGCTGTCATCTCACATGCGATACATGCTTGTCAACCGTTCGTC 1200
Db      1141 GGTTCGGCAGCTCTGCTGTCATCTCACATGCGATACATGCTTGTCAACCGTTCGTC 1200
Qy      1201 TTGTTCCATGTCGCAAGCCCTGCTATTCTGAACCAAGAGATACCTATCCCAAAACAT 1260
Db      1201 TTGTTCCATGTCGCAAGCCCTGCTATTCTGAACCAAGAGATACCTATCCCAAAACAT 1260
Qy      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTGTAAC 1311
Db      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTGTAAC 1311

```

RESULT 3
 US-10-713-381-9
 ; Sequence 9, Application US/10713381
 ; Publication No. US20040221331A1
 ; GENERAL INFORMATION:

```

; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

```

```

Query Match      11.3%; Score 157; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.8e-68;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1155 CGTGTATCTCACATGCGATTAATCACTGTTTCAACCGTTGCTTGTTCATGTC 1214
Db      15  CGTGTATCTCACATGCGATTAATCACTGTTTCAACCGTTGCTTGTTCATGTC 74

```

```

Qy      1215 AAGCTTGGCTTATGTAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 1274
Db      75  AAGCTTGGCTTATGTAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATG 134

```

```

Qy      1275 CAATTCATGCAACACGACATATGTTTCTGTAAC 1311
Db      135 CAATTCATGCAACACGACATATGTTTCTGTAAC 171

```

```

RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

```

```

Query Match      8.0%; Score 111; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 6.9e-45;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1201 TTGTTCCATGTCGCAAGCCCTTCTTGAACCAAGAGATTAATCCCAAAACAT 1260
Db      48  TTGTTCCATGTCGCAAGCCCTTCTTGAACCAAGAGATTAATCCCAAAACAT 107

```

```

Qy      1261 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTGTAAC 1311
Db      108 CCATCTTACTCATGCAACTTCCATGCAACCGCAGATATGTTTCTGTAAC 158

```



```
RESULT 5
US-10-739-930-1857
; Sequence 1857, Application US/10739930
; Publication No. US20040216150A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; PRIOR FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1857
; LENGTH: 1663
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match          5.0%; Score 70; DB 20; Length 1663;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1313 AATCCATTAAAGATCAACAGCTAGCGTTCGCCGCTCCTCTCTCTGCGG 1372
DB      1 AATCCATTAAAGATCAACAGCTAGCGTTCGCCGCTCCTCTCTCTGCGG 60

QY      1373 ATCTTTTCG 1382
DB      61 ATCTTTTCG 70

RESULT 6
US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match          3.6%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1229 CTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATGCAAC 1278
DB      1 CTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATGCAAC 50

RESULT 7
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match          2.9%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1239 AGGATACCTACTCCCAACATTCATCTTACTGATGCAAC 1278
DB      1 AGGATACCTACTCCCAACATTCATCTTACTGATGCAAC 40

RESULT 8
US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match          2.2%; Score 30; DB 20; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1179 CATGCTGTTCACCGTTCGTCCTGTGTCGA 1208
DB      1 CATGCTGTTCACCGTTCGTCCTGTGTCGA 30

RESULT 9
US-09-923-844B-5/c
; Sequence 5, Application US/09923844B
; Patent No. US2002016143A1
; GENERAL INFORMATION:
; APPLICANT: Pioneer Hi-Bred International, Inc.
; APPLICANT: Bao, Zhongmeng
; APPLICANT: Lu, Guihua
; TITLE OF INVENTION: Sclerotinia-inducible Genes and
```



```

; TITLE OF INVENTION: Promoters and Their Uses
; FILE REFERENCE: 35718/234631
; CURRENT APPLICATION NUMBER: US/09/923, 844B
; PRIOR FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/224,603
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 849
; TYPE: DNA
; ORGANISM: Helianthus annuus
US-09-923-844B-5

Query Match          1.6%; Score 23; DB 9; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      854 AGTTATTTTCTTATATAAT 876
DB      556 AGTTATTTTCTTTATAAAT 534

RESULT 10
US-10-085-117-355/C
; Sequence 355, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; PRIOR FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 355
; LENGTH: 194945
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(194945)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-355

Query Match          1.6%; Score 22; DB 17; Length 194945;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      113 TGCATGATCTCCATGTTCCACT 134
DB      28972 TGCATGATCTCCATGTTCCACT 28951

RESULT 11
US-10-357-930-56524
; Sequence 56524, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
```

```

; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56524
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-56524

Query Match          1.5%; Score 21; DB 20; Length 333;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      93 TTACAAAAAAGCTTCCACGGGT 113
DB      89 TTACAAAAAAGCTTCCACGGGT 109

RESULT 12
US-10-425-115-183040
; Sequence 183040, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 183040
; LENGTH: 356
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_98517C.1
US-10-425-115-183040

Query Match          1.5%; Score 21; DB 20; Length 356;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      632 TTTTCTTTTCATCTTGTT 652
DB      72 TTTTCTTTTCATCTTGTT 92

RESULT 13
US-10-425-115-117183/C
; Sequence 117183, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
```



```

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 117183
; LENGTH: 386
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_3835C.1
US-10-425-115-117183

Query Match          1.5%; Score 21; DB 20; Length 386;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 ATGATTTTCGCTTAGGCC 391
DB 367 ATGATTTTCGCTTAGGCC 347

RESULT 14
US-09-864-761-11465/c
; Sequence 11465, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11465
```

```

; LENGTH: 459
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000100.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.2
US-09-864-761-11465

Query Match          1.5%; Score 21; DB 9; Length 459;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1037 AAAGGAGAGAGAGAGAGA 1057
DB 118 AAAGGAGAGAGAGAGAGA 98

RESULT 15
US-10-425-115-51470/c
; Sequence 51470, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 51470
; LENGTH: 1203
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1203)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_146936C.1
US-10-425-115-51470

Query Match          1.5%; Score 21; DB 20; Length 1203;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 848 AAAATTAGTTATTTCTCTT 868
DB 1129 AAAATTAGTTATTTCTCTT 1109
```

Search completed: September 16, 2005, 12:49:09
Job time : 1294.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2005, 10:24:35 ; Search time 4293.5 Seconds
(without alignments)
12358.587 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 cccctggcgtcctcctcgaataaa.....cttttcgcccaccaccacg 1394

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_ests3:*
4: gb_ests4:*
5: gb_ests5:*
6: gb_ests6:*
7: gb_ests7:*
8: gb_ests8:*
9: gb_ests9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	910	65.3	963	9	CC656933 OGDQ20TM
2	745	53.4	915	9	CG224225 OGIAG08TV
3	487	34.9	687	9	CC656939 OGDQ20TV
4	66	4.7	878	9	CG224211 OGIAG08TV
5	27	1.9	539	8	B2828689 B2828689
6	27	1.9	611	8	B2828694 B2828694
7	27	1.9	804	9	CG148339 CG148339
8	26	1.9	854	9	CG011564 CG011564
9	25	1.8	687	8	B2371637 B2371637
10	25	1.8	726	9	CC810403 CC810403
11	25	1.8	819	9	CG305519 CG305519
12	25	1.8	825	9	CG190606 CG190606
13	25	1.8	833	9	CG190609 CG190609
14	25	1.8	857	8	CC358144 CC358144
15	25	1.8	878	9	CG307825 CG307825
16	25	1.8	882	8	CC013382 CC013382
17	25	1.8	892	9	CG197610 CG197610
18	25	1.8	893	8	CC439146 CC439146
19	25	1.8	956	9	CG060573 CG060573
20	24	1.7	385	1	A1626382 A1626382
21	24	1.7	468	1	A1626141 A1626141
22	24	1.7	502	8	BH411902 BH411902
23	24	1.7	533	8	CC010354 CC010354
24	24	1.7	625	9	CG377165 CG377165

25	24	1.7	630	9	CG311717 CG311717
26	24	1.7	639	9	CG009494 CG009494
27	24	1.7	739	9	CG009496 CG009496
28	24	1.7	835	9	CG979335 CG979335
29	24	1.7	857	8	CC35124 CC35124
30	24	1.7	866	9	CG209882 CG209882
31	23	1.6	229	8	CC364841 CC364841
32	23	1.6	318	9	CG421210 CG421210
33	23	1.6	493	1	AL375883 AL375883
34	23	1.6	514	8	B2032120 B2032120
35	23	1.6	533	8	AZ095114 AZ095114
36	23	1.6	610	4	B1676219 B1676219
37	23	1.6	651	8	BH297118 BH297118
38	23	1.6	664	9	CG700282 CG700282
39	23	1.6	691	8	B2524893 B2524893
40	23	1.6	710	8	B2828501 B2828501
41	23	1.6	831	8	B2129958 B2129958
42	23	1.6	833	9	CG437195 CG437195
43	23	1.6	845	9	CG255896 CG255896
44	23	1.6	870	9	CL198382 CL198382
45	23	1.6	900	9	CC708844 CC708844

ALIGNMENTS

RESULT 1	CG656933	963 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGDQ20TM_ZM_0.7_1.5_KB	Zea mays	genomic clone	ZMMA0554D15	
DEFINITION	genomic survey sequence.				
ACCESSION	CC656933				
VERSION	CC656933.1	GI:32060225			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.				
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nuneberg, A., Robbins, D. and Lakey, N.				
TITLE	Unpublished (2002)				
JOURNAL	Consortium for Maize Genomics				
COMMENT	Other GSSs: OGDQ20TV Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends.				
FEATURES	Location/Qualifiers				
source	1..963 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMA0554D15" /clone_1ib="ZM_0.7_1.5_KB" /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	65.3%	Score 910;	DB 9;	Length 963;	
Best Local Similarity	100.0%	Pred. No. 0;			
Matches 910;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
404	ACCAGCAGTGTAGTGTCTATTTGGACATCCAGACAAAGATTTAAAAATAAC	463			
1	ACCAGCAGTGTAGTGTCTATTTGGACATCCAGACAAAGATTTAAAAATAAC	60			


```
QY 464 AAGTAACATACTGAAAGCTATCATGTATGTTTAAAGAAATCTATTTAAACC 523
|
|
|
Db 61 AAGTAACATACTGAAAGCTATCATGTATGTTTAAAGAAATCTATTTAAACC 120
|
|
|
QY 524 AGCATCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTTACAGTTTACAA 583
|
|
|
Db 121 AGCATCTCTTAAACAAAGCATATTTGCAAGAGACAAATATGTTACAGTTTACAA 180
|
|
|
QY 584 CATCTAAGAGCCAAATTTATTCGAAAGGTAGCTATGACGTTTCAGATTTTCTTTTC 643
|
|
|
Db 181 CATCTAAGAGCCAAATTTATTCGAAAGGTAGCTATGACGTTTCAGATTTTCTTTTC 240
|
|
|
QY 644 ATTCTGTATTTTGTATGTTTATATATACATTTTCTCTCTCAATAGAGTATTT 703
|
|
|
Db 241 ATTCTGTATTTTGTATGTTTATATATACATTTTCTCTCTCAATAGAGTATTT 300
|
|
|
QY 704 TCTTCGATTTTATTAATAAGTACTATAAAGTCAATTTTATATTAAGACGATGCTAG 763
|
|
|
Db 301 TCTTCGATTTTATTAATAAGTACTATAAAGTCAATTTTATATTAAGACGATGCTAG 360
|
|
|
QY 764 ATTCTGTCTTAAACATTTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTT 823
|
|
|
Db 361 ATTCTGTCTTAAACATTTTCTGATTTTAAAGAGTATTTGGCAACCTGTTCTTT 420
|
|
|
QY 824 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAA 883
|
|
|
Db 421 CAAGAATTTTGAATTTTCAAAAAAATAGTTTATTTCTCTTATATAAATAGAAAA 480
|
|
|
QY 884 ACTTGAATAATATGAGTGGCCAGACTAGCCCTAGAAATGTTTCCCAATAAATACATCA 943
|
|
|
Db 481 ACTTGAATAATATGAGTGGCCAGACTAGCCCTAGAAATGTTTCCCAATAAATACATCA 540
|
|
|
QY 944 CTGTGTATATTTTGGCCAGGCCCAATAATATTTAAACGAACTGAATTCGAGCA 1003
|
|
|
Db 541 CTGTGTATATTTTGGCCAGGCCCAATAATATTTAAACGAACTGAATTCGAGCA 600
|
|
|
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 1063
|
|
|
Db 601 AACCAATCTGAGCTATTTCTCTAGATTAGTAAAGAGAGAGAGAGAGAGAGAGAGAG 660
|
|
|
QY 1064 TTTTAAAGTCAATGTCCTGAGATGTCGGGTTTGGCAACGATAGCACCGTATCATGCT 1123
|
|
|
Db 661 TTTTAAAGTCAATGTCCTGAGATGTCGGGTTTGGCAACGATAGCACCGTATCATGCT 720
|
|
|
QY 1124 CATAGGTCCTAGCTGAGTTCGGGAGCTCTCGTGCATCTCAATGAGCATATCTACATGC 1183
|
|
|
Db 721 CATAGGTCCTAGCTGAGTTCGGGAGCTCTCGTGCATCTCAATGAGCATATCTACATGC 780
|
|
|
QY 1184 TTGTTCACCGTTCGTCTTGTTCATGTCGCAAGCCCTGCTATTCGAAACCAAGAGAT 1243
|
|
|
Db 781 TTGTTCACCGTTCGTCTTGTTCATGTCGCAAGCCCTGCTATTCGAAACCAAGAGAT 840
|
|
|
QY 1244 ACCTACTCCCAACCAATCTCTTACTCATGCAACTTCGCAACGCAACGCAATATGTT 1303
|
|
|
Db 841 ACCTACTCCCAACCAATCTCTTACTCATGCAACTTCGCAACGCAACGCAATATGTT 900
|
|
|
QY 1304 TCCTGAACCA 1313
|
|
|
Db 901 TCCTGAACCA 910
|
|
|
RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS OGIAG081Y_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0716B15,
DEFINITION genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
1 (bases 1 to 915)
WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
JOURNAL
Consortium for Maize Genomics
Unpublished (2002)
COMMENT
Other GSSs: OGIAG081Y
Contact: Cathy WhiteIaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteIaw@cigr.org
Seq primer: 7F
Classes: sheared ends.
FEATURES
source
location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0716B15"
/clone_1fb="ZM_0.7_1.5_KB"
/clone_2fb="ZM_0.7_1.5_KB"
/clone_3fb="ZM_0.7_1.5_KB"
/clone_4fb="ZM_0.7_1.5_KB"
/clone_5fb="ZM_0.7_1.5_KB"
/clone_6fb="ZM_0.7_1.5_KB"
/clone_7fb="ZM_0.7_1.5_KB"
/clone_8fb="ZM_0.7_1.5_KB"
/clone_9fb="ZM_0.7_1.5_KB"
/clone_10fb="ZM_0.7_1.5_KB"
/clone_11fb="ZM_0.7_1.5_KB"
/clone_12fb="ZM_0.7_1.5_KB"
/clone_13fb="ZM_0.7_1.5_KB"
/clone_14fb="ZM_0.7_1.5_KB"
/clone_15fb="ZM_0.7_1.5_KB"
/clone_16fb="ZM_0.7_1.5_KB"
/clone_17fb="ZM_0.7_1.5_KB"
/clone_18fb="ZM_0.7_1.5_KB"
/clone_19fb="ZM_0.7_1.5_KB"
/clone_20fb="ZM_0.7_1.5_KB"
/clone_21fb="ZM_0.7_1.5_KB"
/clone_22fb="ZM_0.7_1.5_KB"
/clone_23fb="ZM_0.7_1.5_KB"
/clone_24fb="ZM_0.7_1.5_KB"
/clone_25fb="ZM_0.7_1.5_KB"
/clone_26fb="ZM_0.7_1.5_KB"
/clone_27fb="ZM_0.7_1.5_KB"
/clone_28fb="ZM_0.7_1.5_KB"
/clone_29fb="ZM_0.7_1.5_KB"
/clone_30fb="ZM_0.7_1.5_KB"
/clone_31fb="ZM_0.7_1.5_KB"
/clone_32fb="ZM_0.7_1.5_KB"
/clone_33fb="ZM_0.7_1.5_KB"
/clone_34fb="ZM_0.7_1.5_KB"
/clone_35fb="ZM_0.7_1.5_KB"
/clone_36fb="ZM_0.7_1.5_KB"
/clone_37fb="ZM_0.7_1.5_KB"
/clone_38fb="ZM_0.7_1.5_KB"
/clone_39fb="ZM_0.7_1.5_KB"
/clone_40fb="ZM_0.7_1.5_KB"
/clone_41fb="ZM_0.7_1.5_KB"
/clone_42fb="ZM_0.7_1.5_KB"
/clone_43fb="ZM_0.7_1.5_KB"
/clone_44fb="ZM_0.7_1.5_KB"
/clone_45fb="ZM_0.7_1.5_KB"
/clone_46fb="ZM_0.7_1.5_KB"
/clone_47fb="ZM_0.7_1.5_KB"
/clone_48fb="ZM_0.7_1.5_KB"
/clone_49fb="ZM_0.7_1.5_KB"
/clone_50fb="ZM_0.7_1.5_KB"
/clone_51fb="ZM_0.7_1.5_KB"
/clone_52fb="ZM_0.7_1.5_KB"
/clone_53fb="ZM_0.7_1.5_KB"
/clone_54fb="ZM_0.7_1.5_KB"
/clone_55fb="ZM_0.7_1.5_KB"
/clone_56fb="ZM_0.7_1.5_KB"
/clone_57fb="ZM_0.7_1.5_KB"
/clone_58fb="ZM_0.7_1.5_KB"
/clone_59fb="ZM_0.7_1.5_KB"
/clone_60fb="ZM_0.7_1.5_KB"
/clone_61fb="ZM_0.7_1.5_KB"
/clone_62fb="ZM_0.7_1.5_KB"
/clone_63fb="ZM_0.7_1.5_KB"
/clone_64fb="ZM_0.7_1.5_KB"
/clone_65fb="ZM_0.7_1.5_KB"
/clone_66fb="ZM_0.7_1.5_KB"
/clone_67fb="ZM_0.7_1.5_KB"
/clone_68fb="ZM_0.7_1.5_KB"
/clone_69fb="ZM_0.7_1.5_KB"
/clone_70fb="ZM_0.7_1.5_KB"
/clone_71fb="ZM_0.7_1.5_KB"
/clone_72fb="ZM_0.7_1.5_KB"
/clone_73fb="ZM_0.7_1.5_KB"
/clone_74fb="ZM_0.7_1.5_KB"
/clone_75fb="ZM_0.7_1.5_KB"
/clone_76fb="ZM_0.7_1.5_KB"
/clone_77fb="ZM_0.7_1.5_KB"
/clone_78fb="ZM_0.7_1.5_KB"
/clone_79fb="ZM_0.7_1.5_KB"
/clone_80fb="ZM_0.7_1.5_KB"
/clone_81fb="ZM_0.7_1.5_KB"
/clone_82fb="ZM_0.7_1.5_KB"
/clone_83fb="ZM_0.7_1.5_KB"
/clone_84fb="ZM_0.7_1.5_KB"
/clone_85fb="ZM_0.7_1.5_KB"
/clone_86fb="ZM_0.7_1.5_KB"
/clone_87fb="ZM_0.7_1.5_KB"
/clone_88fb="ZM_0.7_1.5_KB"
/clone_89fb="ZM_0.7_1.5_KB"
/clone_90fb="ZM_0.7_1.5_KB"
/clone_91fb="ZM_0.7_1.5_KB"
/clone_92fb="ZM_0.7_1.5_KB"
/clone_93fb="ZM_0.7_1.5_KB"
/clone_94fb="ZM_0.7_1.5_KB"
/clone_95fb="ZM_0.7_1.5_KB"
/clone_96fb="ZM_0.7_1.5_KB"
/clone_97fb="ZM_0.7_1.5_KB"
/clone_98fb="ZM_0.7_1.5_KB"
/clone_99fb="ZM_0.7_1.5_KB"
/clone_100fb="ZM_0.7_1.5_KB"
/clone_101fb="ZM_0.7_1.5_KB"
/clone_102fb="ZM_0.7_1.5_KB"
/clone_103fb="ZM_0.7_1.5_KB"
/clone_104fb="ZM_0.7_1.5_KB"
/clone_105fb="ZM_0.7_1.5_KB"
/clone_106fb="ZM_0.7_1.5_KB"
/clone_107fb="ZM_0.7_1.5_KB"
/clone_108fb="ZM_0.7_1.5_KB"
/clone_109fb="ZM_0.7_1.5_KB"
/clone_110fb="ZM_0.7_1.5_KB"
/clone_111fb="ZM_0.7_1.5_KB"
/clone_112fb="ZM_0.7_1.5_KB"
/clone_113fb="ZM_0.7_1.5_KB"
/clone_114fb="ZM_0.7_1.5_KB"
/clone_115fb="ZM_0.7_1.5_KB"
/clone_116fb="ZM_0.7_1.5_KB"
/clone_117fb="ZM_0.7_1.5_KB"
/clone_118fb="ZM_0.7_1.5_KB"
/clone_119fb="ZM_0.7_1.5_KB"
/clone_120fb="ZM_0.7_1.5_KB"
/clone_121fb="ZM_0.7_1.5_KB"
/clone_122fb="ZM_0.7_1.5_KB"
/clone_123fb="ZM_0.7_1.5_KB"
/clone_124fb="ZM_0.7_1.5_KB"
/clone_125fb="ZM_0.7_1.5_KB"
/clone_126fb="ZM_0.7_1.5_KB"
/clone_127fb="ZM_0.7_1.5_KB"
/clone_128fb="ZM_0.7_1.5_KB"
/clone_129fb="ZM_0.7_1.5_KB"
/clone_130fb="ZM_0.7_1.5_KB"
/clone_131fb="ZM_0.7_1.5_KB"
/clone_132fb="ZM_0.7_1.5_KB"
/clone_133fb="ZM_0.7_1.5_KB"
/clone_134fb="ZM_0.7_1.5_KB"
/clone_135fb="ZM_0.7_1.5_KB"
/clone_136fb="ZM_0.7_1.5_KB"
/clone_137fb="ZM_0.7_1.5_KB"
/clone_138fb="ZM_0.7_1.5_KB"
/clone_139fb="ZM_0.7_1.5_KB"
/clone_140fb="ZM_0.7_1.5_KB"
/clone_141fb="ZM_0.7_1.5_KB"
/clone_142fb="ZM_0.7_1.5_KB"
/clone_143fb="ZM_0.7_1.5_KB"
/clone_144fb="ZM_0.7_1.5_KB"
/clone_145fb="ZM_0.7_1.5_KB"
/clone_146fb="ZM_0.7_1.5_KB"
/clone_147fb="ZM_0.7_1.5_KB"
/clone_148fb="ZM_0.7_1.5_KB"
/clone_149fb="ZM_0.7_1.5_KB"
/clone_150fb="ZM_0.7_1.5_KB"
/clone_151fb="ZM_0.7_1.5_KB"
/clone_152fb="ZM_0.7_1.5_KB"
/clone_153fb="ZM_0.7_1.5_KB"
/clone_154fb="ZM_0.7_1.5_KB"
/clone_155fb="ZM_0.7_1.5_KB"
/clone_156fb="ZM_0.7_1.5_KB"
/clone_157fb="ZM_0.7_1.5_KB"
/clone_158fb="ZM_0.7_1.5_KB"
/clone_159fb="ZM_0.7_1.5_KB"
/clone_160fb="ZM_0.7_1.5_KB"
/clone_161fb="ZM_0.7_1.5_KB"
/clone_162fb="ZM_0.7_1.5_KB"
/clone_163fb="ZM_0.7_1.5_KB"
/clone_164fb="ZM_0.7_1.5_KB"
/clone_165fb="ZM_0.7_1.5_KB"
/clone_166fb="ZM_0.7_1.5_KB"
/clone_167fb="ZM_0.7_1.5_KB"
/clone_168fb="ZM_0.7_1.5_KB"
/clone_169fb="ZM_0.7_1.5_KB"
/clone_170fb="ZM_0.7_1.5_KB"
/clone_171fb="ZM_0.7_1.5_KB"
/clone_172fb="ZM_0.7_1.5_KB"
/clone_173fb="ZM_0.7_1.5_KB"
/clone_174fb="ZM_0.7_1.5_KB"
/clone_175fb="ZM_0.7_1.5_KB"
/clone_176fb="ZM_0.7_1.5_KB"
/clone_177fb="ZM_0.7_1.5_KB"
/clone_178fb="ZM_0.7_1.5_KB"
/clone_179fb="ZM_0.7_1.5_KB"
/clone_180fb="ZM_0.7_1.5_KB"
/clone_181fb="ZM_0.7_1.5_KB"
/clone_182fb="ZM_0.7_1.5_KB"
/clone_183fb="ZM_0.7_1.5_KB"
/clone_184fb="ZM_0.7_1.5_KB"
/clone_185fb="ZM_0.7_1.5_KB"
/clone_186fb="ZM_0.7_1.5_KB"
/clone_187fb="ZM_0.7_1.5_KB"
/clone_188fb="ZM_0.7_1.5_KB"
/clone_189fb="ZM_0.7_1.5_KB"
/clone_190fb="ZM_0.7_1.5_KB"
/clone_191fb="ZM_0.7_1.5_KB"
/clone_192fb="ZM_0.7_1.5_KB"
/clone_193fb="ZM_0.7_1.5_KB"
/clone_194fb="ZM_0.7_1.5_KB"
/clone_195fb="ZM_0.7_1.5_KB"
/clone_196fb="ZM_0.7_1.5_KB"
/clone_197fb="ZM_0.7_1.5_KB"
/clone_198fb="ZM_0.7_1.5_KB"
/clone_199fb="ZM_0.7_1.5_KB"
/clone_200fb="ZM_0.7_1.5_KB"
/clone_201fb="ZM_0.7_1.5_KB"
/clone_202fb="ZM_0.7_1.5_KB"
/clone_203fb="ZM_0.7_1.5_KB"
/clone_204fb="ZM_0.7_1.5_KB"
/clone_205fb="ZM_0.7_1.5_KB"
/clone_206fb="ZM_0.7_1.5_KB"
/clone_207fb="ZM_0.7_1.5_KB"
/clone_208fb="ZM_0.7_1.5_KB"
/clone_209fb="ZM_0.7_1.5_KB"
/clone_210fb="ZM_0.7_1.5_KB"
/clone_211fb="ZM_0.7_1.5_KB"
/clone_212fb="ZM_0.7_1.5_KB"
/clone_213fb="ZM_0.7_1.5_KB"
/clone_214fb="ZM_0.7_1.5_KB"
/clone_215fb="ZM_0.7_1.5_KB"
/clone_216fb="ZM_0.7_1.5_KB"
/clone_217fb="ZM_0.7_1.5_KB"
/clone_218fb="ZM_0.7_1.5_KB"
/clone_219fb="ZM_0.7_1.5_KB"
/clone_220fb="ZM_0.7_1.5_KB"
/clone_221fb="ZM_0.7_1.5_KB"
/clone_222fb="ZM_0.7_1.5_KB"
/clone_223fb="ZM_0.7_1.5_KB"
/clone_224fb="ZM_0.7_1.5_KB"
/clone_225fb="ZM_0.7_1.5_KB"
/clone_226fb="ZM_0.7_1.5_KB"
/clone_227fb="ZM_0.7_1.5_KB"
/clone_228fb="ZM_0.7_1.5_KB"
/clone_229fb="ZM_0.7_1.5_KB"
/clone_230fb="ZM_0.7_1.5_KB"
/clone_231fb="ZM_0.7_1.5_KB"
/clone_232fb="ZM_0.7_1.5_KB"
/clone_233fb="ZM_0.7_1.5_KB"
/clone_234fb="ZM_0.7_1.5_KB"
/clone_235fb="ZM_0.7_1.5_KB"
/clone_236fb="ZM_0.7_1.5_KB"
/clone_237fb="ZM_0.7_1.5_KB"
/clone_238fb="ZM_0.7_1.5_KB"
/clone_239fb="ZM_0.7_1.5_KB"
/clone_240fb="ZM_0.7_1.5_KB"
/clone_241fb="ZM_0.7_1.5_KB"
/clone_242fb="ZM_0.7_1.5_KB"
/clone_243fb="ZM_0.7_1.5_KB"
/clone_244fb="ZM_0.7_1.5_KB"
/clone_245fb="ZM_0.7_1.5_KB"
/clone_246fb="ZM_0.7_1.5_KB"
/clone_247fb="ZM_0.7_1.5_KB"
/clone_248fb="ZM_0.7_1.5_KB"
/clone_249fb="ZM_0.7_1.5_KB"
/clone_250fb="ZM_0.7_1.5_KB"
/clone_251fb="ZM_0.7_1.5_KB"
/clone_252fb="ZM_0.7_1.5_KB"
/clone_253fb="ZM_0.7_1.5_KB"
/clone_254fb="ZM_0.7_1.5_KB"
/clone_255fb="ZM_0.7_1.5_KB"
/clone_256fb="ZM_0.7_1.5_KB"
/clone_257fb="ZM_0.7_1.5_KB"
/clone_258fb="ZM_0.7_1.5_KB"
/clone_259fb="ZM_0.7_1.5_KB"
/clone_260fb="ZM_0.7_1.5_KB"
/clone_261fb="ZM_0.7_1.5_KB"
/clone_262fb="ZM_0.7_1.5_KB"
/clone_263fb="ZM_0.7_1.5_KB"
/clone_264fb="ZM_0.7_1.5_KB"
/clone_265fb="ZM_0.7_1.5_KB"
/clone_266fb="ZM_0.7_1.5_KB"
/clone_267fb="ZM_0.7_1.5_KB"
/clone_268fb="ZM_0.7_1.5_KB"
/clone_269fb="ZM_0.7_1.5_KB"
/clone_270fb="ZM_0.7_1.5_KB"
/clone_271fb="ZM_0.7_1.5_KB"
/clone_272fb="ZM_0.7_1.5_KB"
/clone_273fb="ZM_0.7_1.5_KB"
/clone_274fb="ZM_0.7_1.5_KB"
/clone_275fb="ZM_0.7_1.5_KB"
/clone_276fb="ZM_0.7_1.5_KB"
/clone_277fb="ZM_0.7_1.5_KB"
/clone_278fb="ZM_0.7_1.5_KB"
/clone_279fb="ZM_0.7_1.5_KB"
/clone_280fb="ZM_0.7_1.5_KB"
/clone_281fb="ZM_0.7_1.5_KB"
/clone_282fb="ZM_0.7_1.5_KB"
/clone_283fb="ZM_0.7_1.5_KB"
/clone_284fb="ZM_0.7_1.5_KB"
/clone_285fb="ZM_0.7_1.5_KB"
/clone_286fb="ZM_0.7_1.5_KB"
/clone_287fb="ZM_0.7_1.5_KB"
/clone_288fb="ZM_0.7_1.5_KB"
/clone_289fb="ZM_0.7_1.5_KB"
/clone_290fb="ZM_0.7_1.5_KB"
/clone_291fb="ZM_0.7_1.5_KB"
/clone_292fb="ZM_0.7_1.5_KB"
/clone_293fb="ZM_0.7_1.5_KB"
/clone_294fb="ZM_0.7_1.5_KB"
/clone_295fb="ZM_0.7_1.5_KB"
/clone_296fb="ZM_0.7_1.5_KB"
/clone_297fb="ZM_0.7_1.5_KB"
/clone_298fb="ZM_0.7_1.5_KB"
/clone_299fb="ZM_0.7_1.5_KB"
/clone_300fb="ZM_0.7_1.5_KB"
/clone_301fb="ZM_0.7_1.5_KB"
/clone_302fb="ZM_0.7_1.5_KB"
/clone_303fb="ZM_0.7_1.5_KB"
/clone_304fb="ZM_0.7_1.5_KB"
/clone_305fb="ZM_0.7_1.5_KB"
/clone_306fb="ZM_0.7_1.5_KB"
/clone_307fb="ZM_0.7_1.5_KB"
/clone_308fb="ZM_0.7_1.5_KB"
/clone_309fb="ZM_0.7_1.5_KB"
/clone_310fb="ZM_0.7_1.5_KB"
/clone_311fb="ZM_0.7_1.5_KB"
/clone_312fb="ZM_0.7_1.5_KB"
/clone_313fb="ZM_0.7_1.5_KB"
/clone_314fb="ZM_0.7_1.5_KB"
/clone_315fb="ZM_0.7_1.5_KB"
/clone_316fb="ZM_0.7_1.5_KB"
/clone_317fb="ZM_0.7_1.5_KB"
/clone_318fb="ZM_0.7_1.5_KB"
/clone_319fb="ZM_0.7_1.5_KB"
/clone_320fb="ZM_0.7_1.5_KB"
/clone_321fb="ZM_0.7_1.5_KB"
/clone_322fb="ZM_0.7_1.5_KB"
/clone_323fb="ZM_0.7_1.5_KB"
/clone_324fb="ZM_0.7_1.5_KB"
/clone_325fb="ZM_0.7_1.5_KB"
/clone_326fb="ZM_0.7_1.5_KB"
/clone_327fb="ZM_0.7_1.5_KB"
/clone_328fb="ZM_0.7_1.5_KB"
/clone_329fb="ZM_0.7_1.5_KB"
/clone_330fb="ZM_0.7_1.5_KB"
/clone_331fb="ZM_0.7_1.5_KB"
/clone_332fb="ZM_0.7_1.5_KB"
/clone_333fb="ZM_0.7_1.5_KB"
/clone_334fb="ZM_0.7_1.5_KB"
/clone_335fb="ZM_0.7_1.5_KB"
/clone_336fb="ZM_0.7_1.5_KB"
/clone_337fb="ZM_0.7_1.5_KB"
/clone_338fb="ZM_0.7_1.5_KB"
/clone_339fb="ZM_0.7_1.5_KB"
/clone_340fb="ZM_0.7_1.5_KB"
/clone_341fb="ZM_0.7_1.5_KB"
/clone_342fb="ZM_0.7_1.5_KB"
/clone_343fb="ZM_0.7_1.5_KB"
/clone_344fb="ZM_0.7_1.5_KB"
/clone_345fb="ZM_0.7_1.5_KB"
/clone_346fb="ZM_0.7_1.5_KB"
/clone_347fb="ZM_0.7_1.5_KB"
/clone_348fb="ZM_0.7_1.5_KB"
/clone_349fb="ZM_0.7_1.5_KB"
/clone_350fb="ZM_0.7_1.5_KB"
/clone_351fb="ZM_0.7_1.5_KB"
/clone_352fb="ZM_0.7_1.5_KB"
/clone_353fb="ZM_0.7_1.5_KB"
/clone_354fb="ZM_0.7_1.5_KB"
/clone_355fb="ZM_0.7_1.5_KB"
/clone_356fb="ZM_0.7_1.5_KB"
/clone_357fb="ZM_0.7_1.5_KB"
/clone_358fb="ZM_0.7_1.5_KB"
/clone_359fb="ZM_0.7_1.5_KB"
/clone_360fb="ZM_0.7_1.5_KB"
/clone_361fb="ZM_0.7_1.5_KB"
/clone_362fb="ZM_0.7_1.5_KB"
/clone_363fb="ZM_0.7_1.5_KB"
/clone_364fb="ZM_0.7_1.5_KB"
/clone_365fb="ZM_0.7_1.5_KB"
/clone_366fb="ZM_0.7_1.5_KB"
/clone_367fb="ZM_0.7_1.5_KB"
/clone_368fb="ZM_0.7_1.5_KB"
/clone_369fb="ZM_0.7_1.5_KB"
/clone_370fb="ZM_0.7_1.5_KB"
/clone_371fb="ZM_0.7_1.5_KB"
/clone_372fb="ZM_0.7_1.5_KB"
/clone_373fb="ZM_0.7_1.5_KB"
/clone_374fb="ZM_0.7_1.5_KB"
/clone_375fb="ZM_0.7_1.5_KB"
/clone_376fb="ZM_0.7_1.5_KB"
/clone_377fb="ZM_0.7_1.5_KB"
/clone_378fb="ZM_0.7_1.5_KB"
/clone_379fb="ZM_0.7_1.5_KB"
/clone_380fb="ZM_0.7_1.5_KB"
/clone_381fb="ZM_0.7_1.5_KB"
/clone_382fb="ZM_0.7_1.5_KB"
/clone_383fb="ZM_0.7_1.5_KB"
/clone_384fb="ZM_0.7_1.5_KB"
/clone_385fb="ZM_0.7_1.5_KB"
/clone_386fb="ZM_0.7_1.5_KB"
/clone_387fb="ZM_0.7_1.5_KB"
/clone_388fb="ZM_0.7_1.5_KB"
/clone_389fb="ZM_0.7_1.5_KB"
/clone_390fb="ZM_0.7_1.5_KB"
/clone_391fb="ZM_0.7_1.5_KB"
/clone_392fb="ZM_0.7_1.5_KB"
/clone_393fb="ZM_0.7_1.5_KB"
/clone_394fb="ZM_0.7_1.5_KB"
/clone_395fb="ZM_0.7_1.5_KB"
/clone_396fb="ZM_0.7_1.5_KB"
/clone_397fb="ZM_0.7_1.5_KB"
/clone_398fb="ZM_0.7_1.5_KB"
/clone_399fb="ZM_0.7_1.5_KB"
/clone_400fb="ZM_0.7_1.5_KB"
/clone_401fb="ZM_0.7_1.5_KB"
/clone_402fb="ZM_0.7_1.5_KB"
/clone_403fb="ZM_0.7_1.5_KB"
/clone_404fb="ZM_0.7_1.5_KB"
/clone_405fb="ZM_0.7_1.5_KB"
/clone_406fb="ZM_0.7_1.5_KB"
/clone_407fb="ZM_0.7_1.5_KB"
/clone_408fb="ZM_0.7_1.5_KB"
/clone_409fb="ZM_0.7_1.5_KB"
/clone_410fb="ZM_0.7_1.5_KB"
/clone_411fb="ZM_0.7_1.5_KB"
/clone_412fb="ZM_0.7_1.5_KB"
/clone_413fb="ZM_0.7_1.5_KB"
/clone_414fb="ZM_0.7_1.5_KB"
/clone_415fb="ZM_0.7_1.5_KB"
/clone_416fb="ZM_0.7_1.5_KB"
/clone_417fb="ZM_0.7_1.5_KB"
/clone_418fb="ZM_0.7_1.5_KB"
/clone_419fb="ZM_0.7_1.5_KB"
/clone_420fb="ZM_0.7_1.5_KB"
/clone_421fb="ZM_0.7_1.5_KB"
/clone_422fb="ZM_0.7_1.5_KB"
/clone_423fb="ZM_0.7_1.5_KB"
/clone_424fb="ZM_0.7_1.5_KB"
/clone_425fb="ZM_0.7_1.5_KB"
/clone_426fb="ZM_0.7_1.5_KB"
/clone_427fb="ZM_0.7_1.5_KB"
/clone_428fb="ZM_0.7_1.5_KB"
/clone_429fb="ZM_0.7_1.5_KB"
/clone_430fb="ZM_0.7_1.5_KB"
/clone_431fb="ZM_0.7_1.5_KB"
/clone_432fb="ZM_0.7_1.5_KB"
/clone_433fb="ZM_0.7_1.5_KB"
/clone_434fb="ZM_0.7_1.5_KB"
/clone_435fb="ZM_0.7_1.5_KB"
/clone_436fb="ZM_0.7_1.5_KB"
/clone_437fb="ZM_0.7_1.5_KB"
/clone_438fb="ZM_0.7_1.5_KB"
/clone_439fb="ZM_0.7_1.5_KB"
/clone_440fb="ZM_0.7_1.5_KB"
/clone_441fb="ZM_0.7_1.5_KB"
/clone_442fb="ZM_0.7_1.5_KB"
/clone_443fb="ZM_0.7_1.5_KB"
/clone_444fb="ZM_0.7_1.5_KB"
/clone_445fb="ZM_0.7_1.5_KB"
/clone_446fb="ZM_0.7_1.5_KB"
/clone_447fb="ZM_0.7_1.5_KB"
/clone_448fb="ZM_0.7_1.5_KB"
/clone_449fb="ZM_0.7_1.5_KB"
/clone_450fb="ZM_0.7_1.5_KB"
/clone_451fb="ZM_0.7_1.5_KB"
/clone_452fb="ZM_0.7_1.5_KB"
/clone_453fb="ZM_0.7_1.5_KB"
/clone_454fb="ZM_0.7_1.5_KB"
/clone_455fb="ZM_0.7_1.5_KB"
/clone_456fb="ZM_0.7_1.5_KB"
/clone_457fb="ZM_0.7_1.5_KB"
/clone_458fb="ZM_0.7_1.5_KB"
/clone_459fb="ZM_0.7_1.5_KB"
/clone_460fb="ZM_0.7_1.5_KB"
/clone_461fb="ZM_0.7_1.5_KB"
/clone_462fb="ZM_0.7_1.5_KB"
/clone_463fb="ZM_0.7_1.5_KB"
/clone_464fb="ZM_0.7_1.5_KB"
/clone_465fb="ZM_0.7_1.5_KB"
/clone_466fb="ZM_0.7_1.5_KB"
/clone_467fb="ZM_0.7_1.5_KB"
/clone_468fb="ZM_0.7_1.5_KB"
/clone_469fb="ZM_0.7_1.5_KB"
/clone_470fb="ZM_0.7_1.5_KB"
/clone_471fb="ZM_0.7_1.5_KB"
/clone_472fb="ZM_0.7_1.5_KB"
/clone_473fb="ZM_0.7_1.5_KB"
/clone_474fb="ZM_0.7_1.5_KB"
/clone_475fb="ZM_0.7_1.5_KB"
/clone_476fb="ZM_0.7_1.5_KB"
/clone_477fb="ZM_0.7_1.5_KB"
/clone_478fb="ZM_0.7_1.5_KB"
/clone_479fb="ZM_0.7_1.5_KB
```


Oy	1245	CCTATCTCCCAACCAATCCATCTTACTCATGCAACTTCATGCAACAGCAACATATGTT	1304
Db	624	CCTACTCCCAACCAATCCATCTTACTCATGCAACTTCATGCAACAGCAACATATGTT	683
Oy	1305	CCTGAACCAATTCATTAAAGATCACAAGAGTAGCGTTCTCCGCTAGCTTCCCTCTCTC	1364
Db	684	CCTGAACCAATTCATTAAAGATCACAAGAGTAGCGTTCTCCGCTAGCTTCCCTCTCTC	743
Oy	1365	CTCTGCCGATCTTTTTCGTCCACCA	1389
Db	744	CTCTGCCGATCTTTTTCGTCCACCA	768
RESULT 3			
LOCUS	CC656939/c	687 bp	DNA linear
DEFINITION	OGMDQ20TV_ZM_0.7.1.5_KB_Zea_mays genomic clone ZMMBMA0554D15,		
ACCESSION	CC656939		
VERSION	CC656939.1	GI:32060231	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 687) WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Reenick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Numborg,A., Robbins,D. and Lakey,N. Consortium for Maize Genomics Unpublished (2002) Other_GSSs: OGMDQ20TM Contact: Cathy WhiteJaw TIGR		
TITLE	9712 Medical Center Drive, Rockville, MD 20850, USA		
JOURNAL	Tel: 301-838-5843		
COMMENT	Fax: 301-838-0208		
	Email: whiteJaw@tigr.org		
	Seq primer: TP		
	Class: sheared ends.		
FEATURES	Location/Qualifiers		
source	1..687		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:4577"		
	/clone_1fb="ZMMBMA0554D15"		
	/clone_1lb="ZM_0.7.1.5_KB"		
	/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb		
	methylation filtered genomic DNA library"		
ORIGIN			
Query Match	34.9%; Score 487; DB 9; Length 687;		
Best Local Similarity	100.0%; Pred. No. 1.2e-237;		
Matches	487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Oy	903	CCAGACTAGCCCTAGATGTTTTCCTCAATAATTACATCACTGTGTATTAATTGGC	962
Db	673	CCAGACTAGCCCTAGATGTTTTCCTCAATAATTACATCACTGTGTATTAATTGGC	614
Oy	963	CAGCCCCATAATTATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCGAGCTATTT	1022
Db	613	CAGCCCCATAATTATTTTAAACCGAAATCGAAATCGAGCGAAACCAATCGAGCTATTT	554
Oy	1023	CTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTTGCCCTG	1082
Db	553	CTTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAGTCACTTGCCCTG	494
Oy	1083	AGATGTGGGTTTGGCAACGATAGCCACCGTATCATATAGCTCATATAGTCTTACGTCAAG	1142
Db	493	AGATGTGGGTTTGGCAACGATAGCCACCGTATCATATAGTCTTACGTCAAGTCAAG	434

QY	1145	TTGGGAGAGCTCGTGTGCATCTCA	CAAGGGAATCTAATGCTTGTTCACCGTTCGCTT	1202
Db	433	TTGGGAGAGCTTCGTGTGCATCTCA <td>CAAGGGAATCTAATGCTTGTTCACCGTTCGCTT</td> <td>374</td>	CAAGGGAATCTAATGCTTGTTCACCGTTCGCTT	374
QY	1203	GTTCATCTGTCACCAAGCGCTTCCTATT	TGTGAACCAAGAGATPACTACTCCAAACAATCC	1262
Db	373	GTTCATCTGTCACCAAGCGCTTCCTATT	TGTGAACCAAGAGATPACTACTCCAAACAATCC	314
QY	1263	ATCTTACTCATGCAACTTCATGCAAA	CAGCACATATGTTTCTGTAACCAATCATTTAA	1322
Db	313	ATCTTACTCATGCAACTTCATGCAAA	CAGCACATATGTTTCTGTAACCAATCATTTAA	254
QY	1232	AGATCACAACAGCTAGGGGTTCTCCCG	TAGCTTCCTCCCTCCCTCCGATCTTTTCG	1382
Db	253	AGATCACAACAGCTAGGGGTTCTCCCG	TAGCTTCCTCCCTCCCTCCGATCTTTTCG	194
QY	1383	TCACCA	1389	
Db	193	TCACCA	187	
RESULT 4				
LOCUS	CG224211/c	878 bp	DNA	linear
DEFINITION	OGIAG08TH_ZM_0.7_1.5_KB	Zea mays	genomic clone	ZMMBMA0716B15,
ACCESSION	CG224211			
VERSION	CG224211.1	GI:34124099		
KEYWORDS	GS8.			
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 878) WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Utecherback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlffing,T., Cletek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.			
TITLE	Consortium for Maize Genomics			
JOURNAL	Unpublished (2002)			
COMMENT	Other GS8s: OGIAG08TV Contact: Cathy WhiteLaw			
FEATURES				
source	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@tigr.org Seq primer: TR Class: sheared ends.			
	Location/Qualifiers			
	1..878			
	/organism="Zea mays"			
	/mol_type="genomic DNA"			
	/strain="B73"			
	/db_xref="taxon:4577"			
	/clone_lib="ZMMBMA0716B15"			
	/clone_1lb="ZM_0.7_1.5_KB"			
	/note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"			
ORIGIN				
Query Match	4.7%; Score 66; DB 9; Length 878;			
Best Local Similarity	100.0%; Pred. No. 5.4e-22;			
Matches	66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1324	GATCACAACAGCTAGCGTTCTCCCGCTAGGCTTCCTCTCTGCGCGATCTTTTTCGT	1383	
Db	878	GATCACAACAGCTAGCGTTCTCCCGCTAGGCTTCCTCTCTGCGCGATCTTTTTCGT	819	
QY	1384	CCACCA	1389	
Db	818	CCACCA	813	

RESULT 5
LOCUS BZ828689 539 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFHM74TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA320N03,
genomic survey sequence.
ACCESSION BZ828689
VERSION BZ828689.1 GI:29050496
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 539)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUFHM74TD
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..539
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMMBTA320N03"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 539;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 843 CAAAAAATTAGTTATTTCTCTTT 869
|||||
DB 292 CAAAAAATTAGTTATTTCTCTTT 266

RESULT 6
LOCUS BZ828694 611 bp DNA linear GSS 18-MAR-2003
DEFINITION PUFHM74TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA320N03,
genomic survey sequence.
ACCESSION BZ828694
VERSION BZ828694.1 GI:29050508
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 611)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUFHM74TB
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..611
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMMBTA320N03"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 8; Length 611;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 843 CAAAAAATTAGTTATTTCTCTTT 869
|||||
DB 222 CAAAAAATTAGTTATTTCTCTTT 248

RESULT 7
LOCUS CG148339 804 bp DNA linear GSS 21-AUG-2003
DEFINITION PUKB646TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTA0780G19,
genomic survey sequence.
ACCESSION CG148339
VERSION CG148339.1 GI:34039122
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 804)
Whiteaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUKB646TD
Contact: Cathy Whiteaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteaw@cigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..804
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZMMBTA0780G19"
/clone_1lb="ZM_0.6_1.0_KB"
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"

ORIGIN
Query Match 1.9%; Score 27; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0.052;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 837 TTTTTCACAAAAATTAGTTATTTT 863
|||||
DB 156 TTTTTCACAAAAATTAGTTATTTT 182

RESULT 8
CG011564/c 854 bp DNA linear GSS 19-AUG-2003
DEFINITION ZUABT23TV ZM 3.0 4.0 KB Zea mays genomic clone ZWMBPa0018C22,
genomic survey sequence.
ACCESSION CG011564
VERSION CG011564.1 GI:33883730
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 854)
AUTHORS White, C.A., Quackenbush, J., Van Aken, S., Utecher, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
COMMENT Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: ZUABT23TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@cigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
source location/Qualifiers
1..854
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_1lb="ZM3.0 4.0 KB"
/note="Vector: pBCK-; Site_1: HincII; 3-4 kb 'unfiltered'
genomic DNA library"

ORIGIN
Query Match 1.9%; Score 26; DB 9; Length 854;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 26; Conservative 0; Mismatches 0;

QY 268 GGGGATATATGTCCTCCACATGTC 293
Db 221 GGGGATATATGTCCTCCACATGTC 196

RESULT 9
BZ371637/c 687 bp DNA linear GSS 26-NOV-2002
LOCUS 1e38n03.5', genomic survey sequence.
DEFINITION 1e38n03.5', genomic survey sequence.
ACCESSION BZ371637
VERSION BZ371637.1 GI:25455537
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
AUTHORS Rabinowicz, P.D., O'Shaughnessy, A.L., Ballia, V., Dedhia, N.,
Katzenburger, F., King, L., Miller, B., Muller, S., Nasimimento, L.,
Zutavern, T., McCombie, W.R. and Martienssen, R.A.
COMMENT Genomic shotgun sequences from Zea mays (methyl-filtered)
Unpublished (2002)
Contact: W. Richard McCombie
Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mccombie@cshl.org
Plate: 1e38 row: h column: 03
Seq primer: -21M13univFwd
Class: shotgun
High quality sequence stop: 687.

FEATURES
source location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_1lb="1e38n03"
/lab_host="DH5a"
/clone_1lb="WGS-Zmaysr (DH5a methyl filtered)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
The vector was digested with XbaI and one nucleotide was
added by fill in in the recessive 3' end. The genomic DNA
was nebulized, end repaired, adaptor ligated and size
fractionated using sephadex. The resulting fragments were
between 0.8 and 3 kb and were cloned into the vector (x/y
reads in M13mp19, b/g reads in pUC19). The same ligation
was transformed into DH5a."

ORIGIN
Query Match 1.8%; Score 25; DB 8; Length 687;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 845 AAAAAATTAGTTTATTTCTCTT 869
Db 363 AAAAAATTAGTTTATTTCTCTT 339

RESULT 10
CC810403/c 726 bp DNA linear GSS 16-JUL-2003
LOCUS ZWMBc0479C04r ZWMBc Zea mays genomic clone ZWMBc0479C04 3',
genomic survey sequence.
DEFINITION ZWMBc0479C04r ZWMBc Zea mays genomic clone ZWMBc0479C04 3',
genomic survey sequence.
ACCESSION CC810403
VERSION CC810403.1 GI:32817028
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 726)
AUTHORS Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
Rouzaud, K., Fuks, G., Yu, Y., Wang, R. and Messing, J.
COMMENT Sequencing of the maize genome at PGIR (2003b)
Unpublished (2003)
Contact: Bharti, A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 397.

FEATURES
source location/Qualifiers
1..726
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_1lb="ZWMBc0479C04"
/lab_host="E. coli DH10B"


```

/clone.lib="ZMMB8B"
/notes=Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match 1.8%; Score 25; DB 9; Length 726;
Beet Local Similarity 100.0%; Pred.No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 265 CCAGGGATATATGTCCTCCACCAAT 269
|||||
|||||

Db 364 CCAGGGGATATATGTCCTCCACCAAT 340

RESULT 11
CG305519/c
LOCUS 819 bp DNA linear GSS 25-AUG-2003
DEFINITION OG06GM20TV_ZM_0.7.1.5_KB_Zea_mays genomic clone ZMMBma0713C15,
genomic survey sequence.
ACCESSION CG305519
VERSION CG305519.1 GI:34219733
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citeck,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OG06GM20TH
Contact: Cathy Whiteclaw
TIGR

TITLE 9712 Medical Center Drive, Rockville, MD 20850, USA
JOURNAL Tel: 301-838-5843
COMMENT Fax: 301-838-0208
Email: whiteclaw@cigr.org
Seq primer: 7P
Classes: sheared ends.
Location/Qualifiers
1..819
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone.lib="ZMMBma0713C15"
/clone.lib="ZM_0.7.1.5_KB"
/notes=Vector: pBSCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

FEATURES
Source
1..819
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone.lib="ZMMBma0713C15"
/clone.lib="ZM_0.7.1.5_KB"
/notes=Vector: pBSCSK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 1.8%; Score 25; DB 9; Length 819;
Beet Local Similarity 100.0%; Pred.No. 0.55;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 AAAAAAATTACTTATTTTCTCTT 868
|||||
|||||

Db 142 AAAAAAATTAGTTTATTTCTCTT 118

RESULT 12
CG190606/c
LOCUS 825 bp DNA linear GSS 21-AUG-2003
DEFINITION PUH982TB_ZM_0.6.1.0_KB_Zea_mays genomic clone ZMMBTR0585M20,
genomic survey sequence.
ACCESSION CG190606
VERSION CG190606.1 GI:34081667
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

REFERENCE	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	1 (bases 1 to 825) WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J.			
TITLE	Maize Genomics Consortium			
JOURNAL	Unpublished (2003)			
COMMENT	Other GSSs: PUTHP82TD Contact: Cathy WhiteLaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@cigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers 1..825 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBta0585M20" /clone_1b="ZM_0.6_1.0_KB" /note="Vector: PCR4-T0F0; Site 1: EcoRI; 0.6-1.0 kb high Cor detected genomic DNA library"			
ORIGIN				
Query Match	1.8%;	Score 25;	DB 9;	Length 825;
Best Local Similarity	100.0%;	Pred. No. 0.55;		
Matches	25;	Conservative	0;	Mismatches 0;
Indels	0;	Gaps	0;	
Oy	845 AAAAAATTAGTTATTTCTCTTT 869 715 AAAAAATTAGTTATTTCTCTTT 691			
Db				
RESULT 13				
LOCUS	CG190609 833 bp DNA linear GSS 21-AUG-2003			
DEFINITION	PUIHP82TD ZM_0.6_1.0_KB Zea mays genomic clone ZMMBta0585M20,			
ACCESSION	CG190609			
VERSION	CG190609.1 GI:34081670			
KEYWORDS	GSS.			
SOURCE	Zea mays			
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 833) WhiteLaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezen,J. Maize Genomics Consortium Unpublished (2003) Other GSSs: PUTHP82TB Contact: Cathy WhiteLaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@cigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers 1..833 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBta0585M20"			
FEATURES				
SOURCE				

ORIGIN
 /clone lib="ZM_0.6-1.0_KB"
 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 Cor selected genomic DNA library"

Query Match 1.8%; Score 25; DB 9; Length 833;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 845 AAAAAAATTAGTTATTTCTCTT 869
 |||||
 Db 582 AAAAAAATTAGTTATTTCTCTT 606

RESULT 14
 CC358144/c 857 bp DNA linear GSS 16-MAY-2003
 LOCUS
 DEFINITION
 CC358144
 PUHPY12TD ZM_0.6-1.0_KB Zea mays genomic clone ZMMBtra507B23,
 genomic survey sequence.
 ACCESSION
 CC358144 GI:30827544
 VERSION
 CC358144.1
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays

REFERENCE
 AUTHORS
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 857)
 White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Resnick, A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Bennett, J., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Bennett, J.

TITLE
 JOURNAL
 COMMENT
 Unpublished (2003)
 Other GSSs: PUHPY12TB
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES
 source
 location/Qualifiers
 1..857
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBtra507B23"
 /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
 Cor selected genomic DNA library"

ORIGIN

Query Match 1.8%; Score 25; DB 8; Length 857;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 845 AAAAAAATTAGTTATTTCTCTT 869
 |||||
 Db 853 AAAAAAATTAGTTATTTCTCTT 829

RESULT 15
 CG307825 878 bp DNA linear GSS 26-AUG-2003
 LOCUS
 DEFINITION
 CG307825
 OGMBO11TH ZM_0.7-1.5_KB Zea mays genomic clone ZMMBma0527A01,
 genomic survey sequence.
 ACCESSION
 CG307825 GI:34224985
 VERSION
 CG307825.1
 KEYWORDS
 GSS.
 SOURCE
 Zea mays
 ORGANISM
 Zea mays

REFERENCE
 AUTHORS
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 878)
 White, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
 Resnick, A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Bennett, J., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Bennett, J.

TITLE
 JOURNAL
 COMMENT
 Unpublished (2002)
 Other GSSs: OGMBO11TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source
 location/Qualifiers
 1..878
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMMBma0527A01"
 /note="Vector: pCR4-TOPO; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 878;
 Best Local Similarity 100.0%; Pred. No. 0.55;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 844 AAAAAAATTAGTTATTTCTCTT 868
 |||||
 Db 167 AAAAAAATTAGTTATTTCTCTT 143

Search completed: September 16, 2005, 17:48:56
 Job time : 4301.5 secs

This Page Blank (uspto)

181 CTGACCGAGGCCCATCAGACACCTTTCGGGACACCCATCAAGGCCCTTTCGGATGCCCA 240

Db	181	CTGACCGGAGGCCCATCAGACACCTTTGGGACACCCATCAGAGGCGCTTTCGATGGCCCA	240
Oy	241	CGAGACGATCGGGCGTGTGTATCAGGGGATATATGTCCCCCAATTCGTCACTATA	300
Db	241	CGAGACGATCGGGCGTGTGTATCAGGGGATATATGTCCCCCAATTCGTCACTATA	300
Oy	301	TTATTTATCTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT	360
Db	301	TTATTTATCTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGTATTTTAAAGT	360
Oy	361	GGGCGCTCAGACATAGATTTTCGTTAGGGCCGAAATGAGAGACAGACATGTCTAGG	420
Db	361	GGGCGCTCAGACATAGATTTTCGTTAGGGCCGAAATGAGAGACAGACATGTCTAGG	420
Oy	421	TCACATATGGGCACTACCCAGAACAGATTTTAAAGTATTTTAAAGTATTTTAAAGT	480
Db	421	TCACATATGGGCACTACCCAGAACAGATTTTAAAGTATTTTAAAGTATTTTAAAGT	480
Oy	481	CGAAAGCTATCATGTATATGTATTAAGAAACATCTATTAAGAACATCTCTTAAAAA	540
Db	481	CGAAAGCTATCATGTATATGTATTAAGAAACATCTATTAAGAACATCTCTTAAAAA	540
Oy	541	CAAGCATATTTGAAAGAGACAAATATATGTACAGTTTACAAACATCTTAAGAGGACAA	600
Db	541	CAAGCATATTTGAAAGAGACAAATATATGTACAGTTTACAAACATCTTAAGAGGACAA	600
Oy	601	TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTCTGTATTTTGT	660
Db	601	TTATATCGAAAGGTAGCTATGACGTTCAAGATTTTCTTTTCATCTCTGTATTTTGT	660
Oy	661	ATTGTTTATATATACATTTTCTCTCTTACAAATAGAGATTTTCTCCGATTTATATA	720
Db	661	ATTGTTTATATATACATTTTCTCTCTTACAAATAGAGATTTTCTCCGATTTATATA	720
Oy	721	ATGACATTAAGTCAATTTTATATAAGACAGCATGTCTAGATTCGGTCAAAATC	780
Db	721	ATGACATTAAGTCAATTTTATATAAGACAGCATGTCTAGATTCGGTCAAAATC	780
Oy	781	TTTCTGATTTTATTAAGAGCTAGTTTGGCAACCCGTTTCTTCAAGAAATTTGATTTT	840
Db	781	TTTCTGATTTTATTAAGAGCTAGTTTGGCAACCCGTTTCTTCAAGAAATTTGATTTT	840
Oy	841	TTTCAAAAAAATTAGTTATTTTCTCTTATTAATAGAAAAACATTGAAAAATAGAGT	900
Db	841	TTCAAAAAAATTAGTTATTTTCTCTTATTAATAGAAAAACATTGAAAAATAGAGT	900
Oy	901	TGCCGACGACGACCCGTAAGATGTTTCCAAATTAATTAACATCAGTGTATATATATG	960
Db	901	TGCCGACGACGACCCGTAAGATGTTTCCAAATTAATTAACATCAGTGTATATATATG	960
Oy	961	GCCAGCCCCCATTAATTTATTAACCCGAAATCGAAGGAAACCAATCTGAGCTAT	1020
Db	961	GCCAGCCCCCATTAATTTATTAACCCGAAATCGAAGGAAACCAATCTGAGCTAT	1020
Oy	1021	TTCTCTAGATTAATTAAGAGAGAGAGAGAGAGAAATCACTTAAATGATCATGTGCC	1080
Db	1021	TTCTCTAGATTAATTAAGAGAGAGAGAGAGAGAAATCACTTAAATGATCATGTGCC	1080
Oy	1081	TGAGATGAGCGGTTTGGACAGATGCCACCGTATCATAGTCTATAGTGCCTACGCTCA	1140
Db	1081	TGAGATGAGCGGTTTGGACAGATGCCACCGTATCATAGTCTATAGTGCCTACGCTCA	1140
Oy	1141	GGTTGCGAGCGCTGTGTCACTCTCAATGGCATCTACATGCTGTCTCAACCGTTGCTG	1200
Db	1141	GGTTGCGAGCGCTGTGTCACTCTCAATGGCATCTACATGCTGTCTCAACCGTTGCTG	1200
Oy	1201	TTGTTTCATCGTCAAGGCTTGCTGCTATTTCTGAACCAAGAGATCACTCCCAACAT	1260
Db	1201	TTGTTTCATCGTCAAGGCTTGCTGCTATTTCTGAACCAAGAGATCACTCCCAACAT	1260
Oy	1261	CCATTTTACTCATGCACTTCCATGCAAAACAGCAGCATATGTCTTCTGAACCATTCAT	1320
Db	1261	CCATTTTACTCATGCACTTCCATGCAAAACAGCAGCATATGTCTTCTGAACCATTCAT	1320

QY	1321	AAAGATCAACAACAGCTAGCGCTTCCTCCCGCTAGAGCTCCCTCTCTCTCTGCGGATCTTTT	1380
Db	1321	AAAGATCAACAACAGCTAGCGCTTCCTCCCGCTAGAGCTCCCTCTCTCTCTGCGGATCTTTT	1380
QY	1381	CGTCACCACCATG 1394	
Db	1381	CGTCACCACCATG 1394	
RESULT 2			
LOCUS	BD062176	1394 bp	DNA linear PAT 27-AUG-2002
DEFINITION	Male tissue-preferred regulatory region and method of using same.		
VERSION	BD062176.1	GI:22607781	
KEYWORDS	JP 2001520523-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1394) Albertsen,M.C., Fox,T.W., Garmaet,C.W., Huffman,G.A. and Kendall,T.L.		
TITLE	Male tissue-preferred regulatory region and method of using same		
JOURNAL	Patent: JP 2001520523-A 1 30-OCT-2001; PIONEER HI BRED INTERNATIONAL INC		
COMMENT	PN JP 2001520523-A/1 PD 30-OCT-2001 PF 19-JUN-1998 JP 199504910 PR 23-JUN-1997 US 08/880499 PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARMAET, GARY A HUFFMAN, PI TIMMY L KENDALL PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC Co7K14/34,C12Q1/68, PC A01H5/00 CC Strandedness: Single; CC Topology: Linear; FH Key Location/Qualifiers.		
FEATURES	source	Location/Qualifiers	
ORIGIN			
Query Match	100.0%;	Score 1394;	DB 6; Length 1394;
Best Local Similarity	100.0%;	Pred. No. 3.7e-270;	Indels 0; Gaps 0;
Matches 1394; Conservative	0; Mismatches		
QY	1	CCATGGGTCTCTATGAAAAGATGATTAACAATGTCTATTCCTGTTTCTTAGGGTCC	60
Db	1	CCATGGGTCTCTATGAAAAGATGATTAACAATGTCTATTCCTGTTTCTTAGGGTCC	60
QY	61	CTTCTTCGCTTATTACTGACTGGAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT	120
Db	61	CTTCTTCGCTTATTACTGACTGGAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT	120
QY	121	CTCATGTTCACATCTTCCCACTCGGGGTGGACATTTCTTGGATGTGGTGGTCCCAT	180
Db	121	CTCATGTTCACATCTTCCCACTCGGGGTGGACATTTCTTGGATGTGGTGGTCCCAT	180
QY	181	CTGACCGAGGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCTTTGGATGGCCCA	240
Db	181	CTGACCGAGGGCCCATCGACACCTTTGGGACACCCATCAAGGGCCTTTGGATGGCCCA	240
QY	241	CGAGACGTATCGGGTCTGTGTGATCCAGGGGATATATGTCCCCACAATGTCACCTATA	300
Db	241	CGAGACGTATCGGGTCTGTGTGATCCAGGGGATATATGTCCCCACAATGTCACCTATA	300
QY	301	TTATTTATCTTTAGATATTTATTTATTTTGGAAAAATACAAACTTATCTTTTGTGTA	360
Db	301	TTATTTATCTTTAGATATTTATTTATTTTGGAAAAATACAAACTTATCTTTTGTGTA	360

Db 241 CGAGAGCTATCGG9GTG6NGATCCAGGGGATATATGTCCCCACAATGTCACCTATTA 300
Qy 301 TTATATCTTTAGATATTTATATTTTGGAAAAATACAAACTTATCTTTGTGTA 360
Db 301 TTATATCTTTAGATATTTATATTTTGGAAAAATACAAACTTATCTTTGTGTA 360
Qy 361 GGGCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGACATGCTAGTG 420
Db 361 GGGCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGACATGCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Qy 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTTAAACCAAGATCCTCTTAAAAAA 540
Qy 541 CAAGCATATTTCCAAAAGAGACAATTTGTTAAGTTTACAAACATCTAAGAGCGACAA 600
Db 541 CAAGCATATTTCCAAAAGAGACAATTTGTTAAGTTTACAAACATCTAAGAGCGACAA 600
Qy 601 TTATATCGAAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTTAATTTGTT 660
Db 601 TTATATCGAAAAGGTAAAGCTATGACGTTCAAGATTTTCTTTCAATCTTGTTAATTTGTT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTCTTAATAGATGATTTTCTTCCGATTTTATATA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTCTTAATAGATGATTTTCTTCCGATTTTATATA 720
Qy 721 ATGACTATAAAGTCAATTTTATATTAAGAGACGACATGCTAGATTTCTCGTCAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATTAAGAGACGACATGCTAGATTTCTCGTCAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCCCTGTTCTTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCCCTGTTCTTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTTATTTTCTCTTTATAAAAATGAAAACACTTAGAAAAATGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTTCTCTTTATAAAAATGAAAACACTTAGAAAAATGAGT 900
Qy 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGATTAATTTTGG 960
Db 901 TGCAGACTAGCCCTAGAAATGTTTTCCCAATAAATTAACAATCACTGATTAATTTTGG 960
Qy 961 GCCAGCCCCATAATTTTAAACCGAAACTGAATGACGGAACCAAACTGAGCTAT 1020
Db 961 GCCAGCCCCATAATTTTAAACCGAAACTGAATGACGGAACCAAACTGAGCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAAGAAATCAGTTTAAAGTCAATGCTCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAAGAAATCAGTTTAAAGTCAATGCTCC 1080
Qy 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAACTATAGTCAATAGTGCTCAAGTCA 1140
Db 1081 TGAGATGTGGGTTTGGCAACGATAGCAACCGTAACTATAGTCAATAGTGCTCAAGTCA 1140
Qy 1141 GGTTCGGCAGCTCTCGTGTATCTCACATGGCATACTACATGCTTGTTCAACCGTTGCTG 1200
Db 1141 GGTTCGGCAGCTCTCGTGTATCTCACATGGCATACTACATGCTTGTTCAACCGTTGCTG 1200
Qy 1201 TTGTTCCATGCTGCAAGCCCTGCTATTTCTGAACCAAGAGATACCTAATCCCAAACTAT 1260
Db 1201 TTGTTCCATGCTGCAAGCCCTGCTATTTCTGAACCAAGAGATACCTAATCCCAAACTAT 1260
Qy 1261 CCATCTTACTATGCAACTTCATGCAACACGCAATATGTTTCTGGAACAATCCAT 1320
Db 1261 CCATCTTACTATGCAACTTCATGCAACACGCAATATGTTTCTGGAACAATCCAT 1320
Qy 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCTGCGATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCTGCGATCTTTT 1380

Qy 1381 CGTCCACCACCATG 1394
Db 1381 CGTCCACCACCATG 1394
RESULT 4
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE
ORGANISM
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1
AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
TITLE Male ribeue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="caxon:4577"
ORIGIN
Query Match 99.7%; Score 1389.2; DB 6; Length 1394;
Best Local Similarity 99.8%; Pred. No. 3.4e-269;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 CCATGAGTGTCTATGAAAAAGATGATCAATGTGTATATCCGTTTCTTAGGGTCC 60
1 CCATGAGTGTCTATGAAAAAGATGATCAATGTGTATATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTTGGCCTTATTAATGATGAAATGGGGTTCAAAAACTTCCAGGGTGCATGAT 120
Db 61 CTTCTTGGCCTTATTAATGATGAAATGGGGTTCAAAAACTTCCAGGGTGCATGAT 120
Qy 121 CTCGATGTTCCACTTCTCCACCTCGCGTTGACATTTCTTGGATGCGTGGTCCCAT 180
Db 121 CTCGATGTTCCACTTCTCCACCTCGCGTTGACATTTCTTGGATGCGTGGTCCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACCTTTCCGGACACCCATCAAGGGCTTTCCGATGGCCA 240
Db 181 CTGACCGAGGCCCATCAGACACCTTTCCGGACACCCATCAAGGGCTTTCCGATGGCCA 240
Qy 241 CGAGACGTATCGGGTCGTGGTGAATCCAGGGGATATATGTCCTCCCAATCGTCACTATA 300
Db 241 CGAGACGTATCGGGTCGTGGTGAATCCAGGGGATATATGTCCTCCCAATCGTCACTATA 300
Qy 301 TTATTAATCTTATAGATATTTATTTTGGAAAAATAACAACTTATCTTTGTGTA 360
Db 301 TTATTAATCTTATAGATATTTATTTTGGAAAAATAACAACTTATCTTTGTGTA 360
Qy 361 GGGCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGACATGCTAGTG 420
Db 361 GGGCCTCAGCATGATTTTGGCTTAGGGCCAGAAAATGCGAGGACCGACATGCTAGTG 420
Qy 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Db 421 TCCACTATTTGGCACTACCCAGAAACAAGATTTAAAAAATAACAAAGTAACTAATCCACT 480
Qy 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTTAAAAACCAAGATCCTCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTTAAAAACCAAGATCCTCTTAAAAAA 540
Qy 541 CAAGCATATTTCCAAAAGAGACAATTTGTTAAGTTTACAAACATCTAAGAGCGACAA 600
Db 541 CAAGCATATTTCCAAAAGAGACAATTTGTTAAGTTTACAAACATCTAAGAGCGACAA 600

QY 601 TTATATCGAAGAGTAAGCTATGAGCTTCAGATTTTCTTTTCACTTCTGTATTTTGT 660
DB 601 TTATATCGAAGAGTAAGCTATGAGCTTCAGATTTTCTTTTCACTTCTGTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTATAA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTATAA 720
QY 721 ATGACTATAAAGTCAATTTTATATAGAGACGCGATGCTGTAGATTTCTCGTCAAAAATC 780
DB 721 ATGACTATAAAGTCAATTTTATATAGAGACGCGATGCTGTAGATTTCTCGTCAAAAATC 780
QY 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAATTTTGTATTT 840
DB 781 TTTCTGATTTTTTTAAGAGCTAGTTTGGCAACCTGTTCTTTCAAGAATTTTGTATTT 840
QY 841 TTCAAAAAATTAAGTATTTTCTTTTATATATAAATGAAAAACACTTGAATAATAGAGT 900
DB 841 TTCAAAAAATTAAGTATTTTCTTTTATATATAAATGAAAAACACTTGAATAATAGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATCAATCACTGTATTAATTTTGG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATCAATCACTGTATTAATTTTGG 960
QY 961 GCCAGCCCCATTAATTTTAAACCGAAATCTGAATGAGCGGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCCATTAATTTTAAACCGAAATCTGAATGAGCGGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTGATTTAGTAAAG 1080
DB 1021 TTCTCTGATTTAGTAAAG 1080
QY 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGTATCATATGCTATGATGCTTACGTCA 1140
DB 1081 TGAAGATGTCGGGTTTGGCAACGATAGCCAGTATCATATGCTATGATGCTTACGTCA 1140
QY 1141 GGTTCGGAGACTCTCGTGTATCTCACATGCGCATCTGCTTGTCAACGGTTCGTC 1200
DB 1141 GGTTCGGAGACTCTCGTGTATCTCACATGCGCATCTGCTTGTCAACGGTTCGTC 1200
QY 1201 TTGTTTCATGCTGTCGAAGCTTGGCTATTTGAAACCAAGAGATACCTACCTCCAAACAT 1260
DB 1201 TTGTTTCATGCTGTCGAAGCTTGGCTATTTGAAACCAAGAGATACCTACCTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCAACTTCATGCAACGACGACATATGTTTCTGAAACCAATCATTT 1320
DB 1261 CCATCTTACTCATGCAACTTCATGCAACGACGACATATGTTTCTGAAACCAATCATTT 1320
QY 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCTGCGATCTTTT 1380
DB 1321 AAAGATCAACAAGCTAGCGTTCTCCGCTAGCTTCCCTCTCTCTCTGCGATCTTTT 1380
QY 1381 CGTCCACCAACGATG 1394
DB 1381 CGTCCACCAACGATG 1394

RESULT 5
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION BD062177.1 GI:22607782
ACCESSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same

JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC C07K14/34, C1201/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 1389.2; DB 6; Length 1394;
Best Local Similarity 99.8%; Pred. No. 3,4e-269;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
DB 1 CCATGCTGCTCTATGAAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY 61 CTTCCTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 CTTCCTCTGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 CTCATGTTTCACCTTCTCCACCTGCGCTTGACATTTCTTGATGATGATGATGATGATGAT 180
DB 121 CTCATGTTTCACCTTCTCCACCTGCGCTTGACATTTCTTGATGATGATGATGATGATGAT 180
QY 181 CTGACCGAGGCCCATATGACACCTTTGCGGACACCCATCAAGGCTTTGCGATGCGCA 240
DB 181 CTGACCGAGGCCCATATGACACCTTTGCGGACACCCATCAAGGCTTTGCGATGCGCA 240
QY 241 CGAGAGATGTCGGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 CGAGAGATGTCGGGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 TTATATCTTTAGATATATATTTTGGAAAAATTAACAACCTTATATCTTTGTGTGA 360
DB 301 TTATATCTTTAGATATATATTTTGGAAAAATTAACAACCTTATATCTTTGTGTGA 360
QY 361 GGGCTCAGCATGATTTTGGCTTGAAGGCGCAAAATGCGAGGACCGCATGCTTGA 420
DB 361 GGGCTCAGCATGATTTTGGCTTGAAGGCGCAAAATGCGAGGACCGCATGCTTGA 420
QY 421 TCCACATTTGGGACATCCAGAACAAAGATTTAAAAAATTAACAAAGTATCAATTCACAT 480
DB 421 TCCACATTTGGGACATCCAGAACAAAGATTTAAAAAATTAACAAAGTATCAATTCACAT 480
QY 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGATCCTTAAAAAA 540
DB 481 CGAAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACCAAGATCCTTAAAAAA 540
QY 541 CAAGCATATTTGAAAGAGACAAATTAATGTAAGTTTCAAGTTTCAAAACATCTAAGGCGCA 600
DB 541 CAAGCATATTTGAAAGAGACAAATTAATGTAAGTTTCAAGTTTCAAAACATCTAAGGCGCA 600
QY 601 TTATATCGAAGAGTAAGCTATGAGCTTCAGATTTTCTTTTCACTTCTGTATTTTGT 660
DB 601 TTATATCGAAGAGTAAGCTATGAGCTTCAGATTTTCTTTTCACTTCTGTATTTTGT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTATAA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTCTTACAAATGAGTGAATTTCTCCGATTTTATAA 720

QY 721 ATGACTATAAGTATTTATATTAAGAGCAGCATGTCAGATTCTCGTCAAAAATC 780
 |||||
 Db 721 ATGACTATAAGTATTTATATTAAGAGCAGCATGTCAGATTCTCGTCAAAAATC 780
 |||||
 QY 781 TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGAATTT 840
 |||||
 Db 781 TTTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGAATTT 840
 |||||
 QY 841 TTTCAAAAAAATTAAGTTATTTTCTCTTATTAATAAATGAAAAACATTAAATAGAT 900
 |||||
 Db 841 TTTCAAAAAAATTAAGTTATTTTCTCTTATTAATAAATGAAAAACATTAAATAGAT 900
 |||||
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTCATCACTGTAATTAATTTG 960
 |||||
 Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCCCAATTAATTCATCACTGTAATTAATTTG 960
 |||||
 QY 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCAGCGAAACCAATCTGAGCTAT 1020
 |||||
 Db 961 GCCAGCCCCATTAATTTTAAACCGAACTGAATTCAGCGAAACCAATCTGAGCTAT 1020
 |||||
 QY 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAGAAATCAGTTTAACTCAATGTCCTC 1080
 |||||
 Db 1021 TTCTCTAGATTAGTAAAGAGAGAGAGAGAAATCAGTTTAACTCAATGTCCTC 1080
 |||||
 QY 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGTGTGCTAGCTCA 1140
 |||||
 Db 1081 TGAGATGTGCGGTTTGGCAAGATAGCCACCGTAATCATAGTGTGCTAGCTCA 1140
 |||||
 QY 1141 GGTTCGGCAGCTCTCGTGCATCTCAGATGGCATACTACATGCTTGTTCACCGTTGCTC 1200
 |||||
 Db 1141 GGTTCGGCAGCTCTCGTGCATCTCAGATGGCATACTACATGCTTGTTCACCGTTGCTC 1200
 |||||
 QY 1201 TTGTTCCATGTCGCAAGCCTTGCTATTCTGAACCAAGAGATACCTACCCCAACAT 1260
 |||||
 Db 1201 TTGTTCCATGTCGCAAGCCTTGCTATTCTGAACCAAGAGATACCTACCCCAACAT 1260
 |||||
 QY 1261 CCATCTTACTCATGCACTTCATGCAACAGCAGCATATGTTTCTGGAACCAATCAT 1320
 |||||
 Db 1261 CCATCTTACTCATGCACTTCATGCAACAGCAGCATATGTTTCTGGAACCAATCAT 1320
 |||||
 QY 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 |||||
 Db 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 |||||
 QY 1381 CGTCCACCAACCATG 1394
 |||||
 Db 1381 CGTCCACCAACCATG 1394
 |||||

RESULT 6
 AX224402
 LOCUS AX224402 255 bp DNA linear PAT 10-SEP-2001
 DEFINITION Sequence 9 from Patent WO016097.
 ACCESSION AX224402
 VERSION AX224402.1 GI:15554644
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1
 REFERENCE
 AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 016097-A 9 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1. .255
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

ORIGIN

Query Match 17.1%; Score 238.8; DB 6; Length 255;
 Best Local Similarity 97.2%; Pred. No. 8.3e-38;
 Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1145 CGGAGCTTCGTGTCATCTCATGCACTAGTCAATGCTTGTTCACCGTTCGTTGT 1204
 |||||
 Db 5 CCGGGATTCGGTCATCTCATGCACTAGTCAATGCTTGTTCACCGTTCGTTGT 64
 |||||
 QY 1205 TCCATGTCGTCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATTCAT 1264
 |||||
 Db 65 TCCATGTCGTCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATTCAT 124
 |||||
 QY 1265 CTTACTCATGCACTTCATGCAACAGCAGCATATGTTTCTGGAACCAATCCATTAAAG 1324
 |||||
 Db 125 CTTACTCATGCACTTCATGCAACAGCAGCATATGTTTCTGGAACCAATCCATTAAAG 184
 |||||
 QY 1325 ATCACAACGATAGCGTTCTCCGCTAGCTTCCTCTCTCTGCGCATCTTTTTCGTC 1384
 |||||
 Db 185 ATCACAACGATAGCGTTCTCCGCTAGCTTCCTCTCTCTGCGCATCTTTTTCGTC 244
 |||||
 QY 1385 CACCAACCATG 1394
 |||||
 Db 245 CACCAACCATG 254
 |||||

RESULT 7
 AX224396
 LOCUS AX224396 158 bp DNA linear PAT 10-SEP-2001
 DEFINITION Sequence 3 from Patent WO016097.
 ACCESSION AX224396
 VERSION AX224396.1 GI:15554638
 KEYWORDS
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1
 REFERENCE
 AUTHORS Albertsen, M.C., Fox, T.W., Garnat, C.W., Huffman, G. and Kendall, T.L.
 TITLE Male tissue-preferred regulatory region and method of using same
 JOURNAL Patent: WO 016097-A 3 23-AUG-2001;
 PIONEER HI-BRED INTERNATIONAL, INC. (US)
 FEATURES
 source
 1. .158
 /organism="Zea mays"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4577"

Query Match 10.5%; Score 146; DB 6; Length 158;
 Best Local Similarity 99.4%; Pred. No. 3.9e-19;
 Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1155 CGTGCATCTCATGCACTAGTCAATGCTTGTTCACCGTTCGTC-TTGTTCATGTC 1213
 |||||
 Db 1 CGTGCATCTCATGCACTAGTCAATGCTTGTTCACCGTTCGTC-TTGTTCATGTC 60
 |||||
 QY 1214 CAAGCTTGCCATTTCTGAACCAAGAGATACCTACTCCCAACCAATCCATTACTCAT 1273
 |||||
 Db 61 CAAGCTTGCCATTTCTGAACCAAGAGATACCTACTCCCAACCAATCCATTACTCAT 120
 |||||
 QY 1274 GCAACTTCATGCAACAGCAGCATATGTTTCTGTAAC 1311
 |||||
 Db 121 GCAACTTCATGCAACAGCAGCATATGTTTCTGTAAC 158
 |||||

RESULT 8
 AC147602
 LOCUS AC147602 186199 bp DNA linear HTG 17-AUG-2004
 DEFINITION Zea mays clone ZMMB0334A01, *** SEQUENCING IN PROGRESS ***. 6
 ordered pieces.
 ACCESSION AC147602
 VERSION AC147602.5 GI:51315585

KEYWORDS
SOURCE
ORGANISM

HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.

Zea mays

Zea mays

REFERENCE

Eukaryote: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.

1 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

2 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

3 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

4 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

5 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

6 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

7 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

8 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

9 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

10 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

11 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

12 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

13 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

14 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

15 (bases 1 to 186199)

Björn, B., Nussbaum, C., Lander, E., Butler, E., Wing, R., Bharti, A.K.

and Messing, J.

Zea mays, clone ZMNBC0334A01

Unpublished

(http://pgir.rutgers.edu)

Butler, E. and Wing, R.: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

----- Project Information

Center project name: L30003

Center clone name: 334_A_1

----- Consensus Information

This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ns are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?l). An exact list of reads used in this assembly are available at http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 100617: contig of 100617 bp in length

* 100618 100717: gap of unknown length

* 100718 104730: contig of 4013 bp in length

* 104731 104830: gap of unknown length

* 104831 115104: contig of 10274 bp in length

* 115105 115204: gap of unknown length

* 115205 115386: contig of 41192 bp in length

* 115387 156496: gap of unknown length

* 156497 179936: contig of 23440 bp in length

* 179937 180036: gap of unknown length

* 180037 180037: contig of 6163 bp in length.

FEATURES

source

1..186199
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMNBC0334A01"
/clone_lib="CHORI-201 Maize B73 BAC MDO1 Library"

ORIGIN

Query Match

Best Local Similarity 71.4%; Pred. No. 2.3e-05;

Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

Db

780 CTTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATT 839

162058 CTTTCTTCTTAACTTTAGGCTAGTTGGCAACCTGATT-TTCTAAAGAAATTTCAATTT 162116

Qy

840 TTTTCAAAAAATTTGTTATTTCTTCTTAAATTAAGAAACCTTGAAGAAATAGAG 899

162117 TTTTAAAGTAAATTTGTTATTTCTTCTTAAAGAAATTAAGAAATTTTGTAAATAGAG 162176

Db

900 TTGCCAGACTAGCCCTAGATGTTTCCCAATTAATTAACA 940

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

(http://pgir.rutgers.edu)

Butler, E. and Wing, R.: Arizona Genomics Institute, Biological Sciences West, 448A, P.O. Box 210088, University of Arizona, Tucson, AZ 85721, USA (http://www.genome.arizona.edu)

----- Project Information

Center project name: L30003

Center clone name: 334_A_1

----- Consensus Information

This consensus is derived from a shotgun assembly that has been manually curated. It is the best representation of the BAC that we can generate without further laboratory work. The draft assembly has been edited, and if possible, ends identified by vector as well as by BAC end sequences, and contigs ordered and oriented. Bases that are not Ns are either above Q20 or manually edited. This assembly was performed with Arachne (Genome Res. 2002 12: 177-189; Genome Res. 2003 13: 91-96). All trace files for this project are available at the NCBI trace repository (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?l). An exact list of reads used in this assembly are available at http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.

* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 100617: contig of 100617 bp in length

* 100618 100717: gap of unknown length

* 100718 104730: contig of 4013 bp in length

* 104731 104830: gap of unknown length

* 104831 115104: contig of 10274 bp in length

* 115105 115204: gap of unknown length

* 115205 115386: contig of 41192 bp in length

* 115387 156496: gap of unknown length

* 156497 179936: contig of 23440 bp in length

* 179937 180036: gap of unknown length

* 180037 180037: contig of 6163 bp in length.

FEATURES

source

1..186199
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMNBC0334A01"
/clone_lib="CHORI-201 Maize B73 BAC MDO1 Library"

ORIGIN

Query Match

Best Local Similarity 71.4%; Pred. No. 2.3e-05;

Matches 115; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

Db

780 CTTTCTGATTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAAATTTGATT 839

162058 CTTTCTTCTTAACTTTAGGCTAGTTGGCAACCTGATT-TTCTAAAGAAATTTCAATTT 162116

Qy

840 TTTTCAAAAAATTTGTTATTTCTTCTTAAATTAAGAAACCTTGAAGAAATAGAG 899

162117 TTTTAAAGTAAATTTGTTATTTCTTCTTAAAGAAATTAAGAAATTTTGTAAATAGAG 162176

Db

900 TTGCCAGACTAGCCCTAGATGTTTCCCAATTAATTAACA 940

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Qy

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

Db

162177 TTTTCAAACTAACCTCAGATATATATGAATGAAGCAACA 162217

ACCESSION AC117267 AC115597
VERSION AC117267.2 GI:42733680
KEYWORDS
SOURCE Dictyostelium discoideum
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
AUTHORS 1 (bases 1 to 25769)
Gloeckner, G., Eichinger, L., Szefranski, K., Pachebat, J., Dear, P.,
Lehmann, R., Baumgart, C., Parra, G., April, J. P., Guigo, R., Kumpf, K.,
Tüngel, B., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A. and
Noegel, A. A.
TITLE Sequence and analysis of chromosome 2 of Dictyostelium discoideum
JOURNAL Nature 418 (6893), 79-85 (2002)
MEDLINE 22092622
PUBMED 12097910
REMARK The Dictyostelium Genome Sequencing Consortium
AUTHORS 2 (bases 1 to 25769)
Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 25769)
Baumgart, C.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2004) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On or before Feb 21, 2004 this sequence version replaced
gi:19570016, gi:20087114.
CDS predictions from Geneid may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.html)
Punding
Agency : Deutsche Forschungsgemeinschaft (DFG).
Location/Qualifiers
1. 25769
/organism="Dictyostelium discoideum"
/mol_type="genomic DNA"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="5836255-5862024"
/complement(join(832..1070,1237..1349,1533..1798,
1896..2053,2231..2360))
/note="Geneid exon scores (in order of location ranges) :
23.68, 1.54, 15.75, 9.21, -4.75 - GSCJ_ID dd_03387"
/codon_start=1
/product="hypotheetical protein"
/protein_id="AAS38625.1"
/db_xref="GI:42733681"
/translation="MRLQILFLPILSLILPOLKSTPSSNDISSTILSTITPNKDVNNO
ILKSEWFRSLKSSNNKVIFFVFDGPNHEIYNLRITIKERENKVIYGVNDSFSCFL
NSKFNLEVIKFKSTISRNKKVYPENIWLNESTHCYNNSVLYNKKNEVEQELV
FSQVAKSFLQNGYNIKISSKNTFKIFENNNDKLFSSLPFYLNKIMENDEFEY
SSPLGWTDFKIKQFKLNDIETILSDVEVNLCCSSNNNNNNNNNNVDETLTLGSSGSG
SGSGSSGSGSGAQLSGTGVYINHO"
join(2957..3173,3272..5583)
/note="Geneid exon scores (in order of location ranges) :
24.88, 254.96 - GSCJ_ID dd_00753"
/codon_start=1
/product="hypotheetical protein"
/protein_id="AAS38626.1"
/db_xref="GI:42733682"
/translation="MEPPTKYVYITSNLDPKRWPFQSEMSLTTTDTIOKSPVNSG
SSSSSKGFGEAVLDLIDTAGQSEYSAMRQVYRTDCFMIVSIDSRSFEEVSLK
QHIERVADRDVPIIIVGNKVDELSRQVRIADQLASRLVPIETIAKRSNIER
AFETLVHTRPNTVYKVVNGGAGSAILIOPINHFVEEDPPIEDSYKROVYIS
GLPIIGSLNKSSSSSSSSSSSGKTGLENKIFSGKDQSPQOAAASGTTDRGCI
STNRLEANTVYSMSNLSEKVLITGDVYCGCCNYLISKQSPQAAASGTFWKEF
CKYSNNTILLEGEIIPKDSVEYVLSPTSSITDSSKKESTIYICIDVSGMGIT
EVSJLSEWNAKGVAGASGPSYISRLCEVSSIPYIMDRISIOYPKRVLVLTFS
DEVMIYTOQSNVSDGPIVIAADKLEDFDQLEIGRSMITYDLPTAAGSSDFLKAKIKSL

EPVOSTALGPALLVSAALASOKMLSEVVICTDGVENVGLAIEDLPLGPAOFEYEKVT
KLAONKTTTINIGSGSHIDLGVIKSGSEONNGTITIHPIELAREIRKLPONMIA
TDVMSICLHPTEINKYDSKQGLSRVYKQFNNVSNISLDTLLYSRRPTEVQOYF
FOIQIKYTLDSVCLRVASAOQLAPDPTNSISANISILMAFLQOAKIAQOEY
MSRHLKAAKATLIRSLCNTDQWEEFNVFEVLRREMEAPLITCIKNQOVRKATD
DEIQVYKMKVHKSFVSGRKKDISRRGAEIRKQYNIKFT"
complement(5711..6796)
/note="Geneid exon scores (in order of location ranges) :
68.61 - GSCJ_ID dd_00752"
/codon_start=1
/product="similar to Plaeomedium falciparum (isolate 3D7).
Hypotheical protein"
/protein_id="AAS38627.1"
/db_xref="GI:42733683"
/translation="MSSRYKSLIERVNGINLTIEKSLSPRGILNTNINNNNNNN
GITTTTSGANNNSSTPSTPVSPFDRVDSNRNQIOIYVDSFPNINNNNNNN
NN
KNNNNNNYFDFSTNEMITTLNLCYLEQIDNCYSTNTPSPSSSSSTSLAN
KSNHMMPSLKINOIINDNNNNININFEVQISNSSSLSSGCTYQSCPIDEHIOE
OLRIQEOLOKIIDDDDDDDDDDDSDSPSPPTTTTKTFCLSRPRIKKEENID
QNSMPPPKLDKRRSRDLIPTANI"
join(8924..9076,9236..10573)
/note="Geneid exon scores (in order of location ranges) :
10.37, 128.60 - GSCJ_ID dd_00729"
/codon_start=1
/product="hypotheical protein"
/protein_id="AAS38628.1"
/db_xref="GI:42733684"
/translation="MSSSTITPIATTTTNNNSSTNGHNSKTNRTLEFKNIKNINPS
PNNAIKNSVGIENGPNVNNNNDRKLSNDGVNGKNLTFPSKNSMSTKHS
SSSSDLAKNEVDIESIKELKSMDSITGSLSPLESSESDISTIDESTATTK
LIDAVAAASSSSSVKTYIDVLSFKSNATKPIIDIVENHOKELVPLSPTPYN
NNNNNNNNSSOHTANQOKYPIESPOISPKFAYAOSSSTNNNNNNNNNTT
TKKANSITTPSNNRNHYNNTSKSKKPKPIPSLNTAIPNOKNTTTPSKSST
TPKSNKINDVNAAPAAVANSITTEQVALIIDATPSTNPKPIKEDTPKRLA
AAQROIDIGMTAGYKNYIALVPKSRKRPDTPNKQOVCKRSMDGQIKKMRQLH
ENYDWTGPIPEKVEVASENSTESTETLEKEMNKLEISENENHNH"
complement(10972..11979)
/note="Geneid exon scores (in order of location ranges) :
23.29 - GSCJ_ID dd_00728"
/codon_start=1
/product="hypotheical protein"
/protein_id="AAS38629.1"
/db_xref="GI:42733685"
/translation="MEKIKFEYFEIIDPONOKFIOBFTSKLIGKESFTTFRPIVY
EYLNWTOSESLIKTNGTINKSIFBIFSCYIGELREYEPRIKFTTPTIIL
DYITERKNSLWMLNEKELINSLIYTNKKENKIGFBWETINEFQNTKESFOR
IINANSMTKSIDSRNISNSPINSRCSFNEFEKRSDDASGFLITLSMIGIMNYI
ENSIIESKEFOTINHAKSPFLINDLVSPNEIENDLANTIKILAIOLANSIOESID
KTIELIIDHYLKFSLSIETILKYQNDSTYQLKQVONSKILSGIYFAHKSKRY
N"
join(12884..13020,13172..15320)
/note="Geneid exon scores (in order of location ranges) :
2.42, 136.56 - GSCJ_ID dd_00727"
/codon_start=1
/product="similar to Dictyostelium discoideum (slime
mold). Hypotheical 97.7 kda protein"
/protein_id="AAS38630.1"
/db_xref="GI:42733686"
/translation="MDNLIYDFLIKEKNQGSFLDCILILSINGLYISYDGLDYCNN
FPLINENKALISYGLANVOYQLDEFTFLKLNKNTDCKRKINTHNSLKKCC
GNKKDYSAILIIEENNNIINKDQFNNKNSNNNIYNNNDNNNSIIEEDLNKII
TLKNERINGDEITTTTTTTSSTTSSTTITMTTGTGNGVNRNLTQVORNVYIK
TOLLFLALYNHANKVFLPFIQVYKFKCYLOSMYLANHNDDDADDESYSDI
DDDSYSDGNCSSSSGSDIGSSSNSINNSISSSSSIIANSLSLSTRVNSL
PEGLASIEFEKENVYIQRLPSPISISIRFSGFQRLAKGVISNVSISTGDSFNO
SLDGKMLKPOLHLOLGHKFOQITMGQLPSLITLIDPRAKGVIEIGSIDPVKT
LIDKNSCSNOSISFNPPIKSIITLVDSSENONIKANDISSNVLISIHGEHNS
DIGISLNSIREIKRGAFPDRLICSSITSIDFGKKRNPPLSMNTQTLSDIFS
KNOQIIPGCIPIHTLKLNLNGRYHNOIIPADLTPYLESLNLGSGNEIYKDNBYD
CYGISKGGFSGSSNSNFCVGGTNGLEBMLKNTISLKTILNTFNKRIEVDLPSI
ESLNLGYHNOQPIGNVPLKLLKFLINSEFNONISADGCIPIGLQTIYIRNSMNF
INSLNPLFTTKYINIIDSLH"

QY	504	AAGAAACATCTATTTAAACACAGATCCCTTTAAAAAACAGATCTTTTGGAAAGACAA	563
Db	621	AATATTTATCTATGATTAATTAAGCATTTGTGGAAATTTATTTTATGGAAATATTAAT	562
QY	564	ATTATGTTACAGTTTACAAACATCTAAGACGCAAAATTATATCGAAAGGTAAAGCTATGA	623
Db	561	ATATATGTAATATTTTATTTTATTAATATATA-AATATATATATATATATATATTTT	503
QY	624	CGTTCAGATTTTCTTTTTCATTTCTGTGTATTTTGTATTTGTTTATATATACATTTTCTT	683
Db	502	TTTAAA	443
QY	684	CTCTTACATATAGAGTATTTTCTCCGATTTTATATAAATGACTATAAGTCATTTTATTA	743
Db	442	TTTTCAGTATATAAAGTATTTTACGGAATATGTGAAATGAAAAAGAAATTTGTATG	383
QY	744	TAAAGCAGCAGATGTCGTATGATCTTCGTTCAAAAATCTTTCGATTTTTTTAAGAGCTAG	803
Db	382	TAAATTTTAAATATAAATAAATAAAGAAATATATTTTTTTTATATAATATAATATAAAG	323
QY	804	TTTGGCAACCCGCTTCTTTCAAGAAATTTGATTTTTCAAAAAATTAAGTTATATTT	863
Db	322	TTTATTTATTTTATATATATATATTTTTTTTTTTTAAATTCATATATTTTTTAAATTT	263
QY	864	CTCTTTATATAAATATAGAAAACTTTAGAAAAATAGAGTTGCCAGATCAGCCCTAGAAATTT	923
Db	262	ATATTTTTTTTTTTTTTTTTTGTCTATATATATATATATATATATATATATATATATATTT	203
QY	924	TTCCCAATTAATTAACATCATCTGTGTATATATTTTGGCCAGCCCCATTAATTTATTA	983
Db	202	TTTTTTTTTTTTTTTTTTTAAATTTGAAATATAATTTATTTTCTCATATATATAATTAATA	143
QY	984	CCGAATCTGAATTCAGCGAAACCAATCTGACTATTTCTCTAGATTAGTAAAAAGGA	1043
Db	142	TAAATTTTATATTTTAAAGGTATCATTAATGTTTTTTTAAAAATAAAAAAAATATGCT	83
QY	1044	GAGAGAGAGAGAAGA 1059	
Db	82	TATCTATATATTAATAA 67	
RESULT 11			
LOCUS	DDU86962/c	3576 bp	DNA linear INV 23-JUL-2001
DEFINITION	Dictyostelium discoideum miga (miga) gene, complete cds.		
ACCESSION	U86962		
VERSION	U86962.1	GI:1841871	
KEYWORDS			
SOURCE	Dictyostelium discoideum		
ORGANISM	Dictyostelium discoideum		
REFERENCE	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.		
AUTHORS	1 (baees 1 to 3576)		
TITLE	Escalante, R., Weesels, D., Soll, D. R. and Loomis, W. F.		
JOURNAL	Chemotaxis to cAMP and slug migration in Dictyostelium both depend		
MEDLINE	on miga, a B7B protein		
PUBMED	Mol. Biol. Cell 8 (9), 1763-1775 (1997)		
REFERENCE	2 (baees 1 to 3576)		
AUTHORS	Escalante, R., Weesels, D., Soll, D. and Loomis, W. F.		
TITLE	Direct Submision		
JOURNAL	Submitted (27-JAN-1997) Department of Biology, University of		
USA	California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-032,		

```

FEATURES
    source
        location/qualifiers
            1..3576
                /organism="Dictyostelium discoideum"
                /mol_type="genomic DNA"
                /strain="AX4"
                /bud_strain="AK244"
                /dd_xref="taxon:44689"
                /chromosome="4"
                /map="4..25"

```

	gene	884.. .3423
	/gene="miga"	
mRNA	Join(884.. .2899,2998.. .>3423)	
	/gene="miga"	
	/product="Miga"	
CDS	Join(884.. .2899,2998.. .3423)	
	/gene="miga"	
	/function="essential for slug migration"	
	/note="N-terminus of this protein is similar to other BRB domain proteins, such as the Mus musculus and human zinc finger proteins encoded by GenBank Accession Numbers U14556 and L16896, respectively"	
	/codon_start=1	
	/product="Miga"	
	/protein_id="BAB47544.1"	
	/db_xref="GI:1841872"	
	/translation="MATINIINNNNNNNNNNNNNNINHTFVFNNGFKSLGFEKKT TDIKCKCGQIIFPTIRILANSSEFFGRLLLSDFESSQSVELKOPDPFNVPWVL FMYDIIIVSPENVPIPLAMSDHYLRDLQHCLFLPDNLRENVLTITAKSIDH EPIRERCTTISKNFTFCENLYKOTRIINN>NNNCINENNKKHLNTSSASSLNS TSSAASSSSSSNNYYTTTTTTTTTTTATSSSFNRGLTNCKRTGCQQQQQQQQQQ QQQQQQQQQQPYDFSFLPASIFYQLLDHDLVVNSEYLRLVALNFLEKFGEENT ITYNNITDOLKSDMSQIRIILMPDYQLTLITDPLPKELPFTLLERKRFEFG TSBSSSSTSSPSMTSNSPTKOTSPSKYQQQQQQQQQQQQQQQQQQQQQQQQ QQQQQQQQQQPOQPQIIOQPPQIIOQPPQIIOQPPQIIOQPPQTICQVQQQQNQ QQQQQQQQQQIQQQQTOMQQOQLQQLIQLOQQQQQQQQQQIIOFTSSSPLESS ISLASOSKLSINNAETDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNKQSSNNNDK LKRIQRPPOSILPEYSGFDPDFKGIIFNISTDGAEKSNPHESTKIKRTSSIDKG LYDIELTPNALFWTKDVPAWSMIDLGNRTVAEMVYTIIRHGLSYKSDSLRWDFQG TNGBWVTILKRTINDPSLNTKYATHSWPTVCETAERYFRILIQTGKNSNNRNFLVG LEIYGELCELTNPBY"	
exon	<884.. .2899	
	/gene="miga"	
	/number=1	
misc_feature	951.. .1289	
	/gene="miga"	
	/note="encodes BTB domain; similar to the Drosophila melanogaster BTB domain encoded by GenBank Accession Number U01333"	
variation	1253^1254	
	/gene="miga"	
	/note="gene interruption in mutant"	
	/replac="pbst external plasmid, approximately 4 kb"	
intron	2900.. .2997	
	/gene="miga"	
	/number=1	
exon	2998.. .3423	
	/gene="miga"	
	/number=2	

Query Match	5.1%	Score 71.2	DB 3	Length 3576
Best Local Similarity	47.2%	Pred. No. 0.0039		
Matches	217	Conservative	0	Matches 243
			Indels	0
			Gaps	0
Qy	412	TGTCAGGTCACCATTTGGCGACTCCAGAAACAAGATTAAAAATAACCAAGTAAC	471	
Db	852	TATTTATTTATTTATTTATTTACATTTTGGTTTGGTATAAAAAATAAAAATAAAAA	793	
Qy	472	TAATCACTCGAAGCTATCATGTATGTTTAAAGAAACATCTATTTAAACACGATCCT	531	
Db	792	TAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAACAACTATTAATTTAA	733	
Qy	532	CTTAAAAACAAGCATTTTCGAAGACAAATATGTTACAGTTTACAACATCTAAG	591	
Db	732	TTTTAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAATATTAATTTAAAAA	673	
Qy	592	AGCGACAATTTATTCGAAGTAAAGCATGCTTCAGATTTTCTTTTCATTCTGT	651	
Db	672	AAAAAAAAAAAAAAAAATATAATATAGGCTGTGGAAGAAATGTTTTTTTTTTGGTTGTG	613	
Qy	652	TATTTGTAATGTTTAAAAATACATTTTCTTCTCTTACAAATAGAGTGATTTTCTTCGA	711	

Db 612 TTGTTATTTTGTGTTATTTAAATTAATTTATTAACATTAATAAACTTTAAACACAC 553

QY 712 TTTTATTAATAATGACTATAAGTCATTTTAATTAAGACGACGATGTCGATTCGT 771

Db 552 ACACACACACACCAATGTTTTTTTTTGTGTTTTTTTGTGTTTTTTTATCAACCT 493

QY 772 TCAAAATCTTTCGATTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAT 831

Db 492 TTTCAATCGCTTTGATTTAGTTTTTTTTTTTTTTTGAATGTTTTTTTGAAGAAAT 433

QY 832 TTTGATTTTTCAAAAAATAGTTATTTCTCTTAT 871

Db 432 ATTTTATTTTATTAATTTTATTTTATTTATTTTATTTT 393

RESULT 12
BX957346 149526 bp DNA linear HTG 08-OCT-2004
LOCUS BX957346.13 GI:54019944
DEFINITION Danio rerio clone CH211-117K16, WORKING DRAFT SEQUENCE.
ACCESSION BX957346
VERSION BX957346.13 GI:54019944
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 149526)
Direct Submission
Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk On Oct 9, 2004 this sequence version replaced gi:53850295.
----- Genome Center.
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc117K16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149526 bases at least Q40
Consensus quality: 149526 bases at least Q30
Consensus quality: 149526 bases at least Q20
Insert size: 149526; sum-of-contigs
Quality coverage: 12.41x in Q20 bases; sum-of-contigs Quality coverage: 12.54x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
source
1. 149526
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-117K16"
/clone_11b="CHORI-211"
1. 149526
/note="assembly_fragment:03115
clone_end:SP6
vector_side:right
clone_end:T7
vector_side:left"

ORIGIN
Query Match 5.1%; Score 70.8; DB 2; Length 149526;
Best Local Similarity 46.8%; Pred. No. 0.00041;

Matches 260; Conservative 0; Mismatches 292; Indels 4; Gaps 1;

QY 488 TATCATGTATGTTTAAAGAAACATCTATTAACACGATCCTTAACAAACAGCAT 547

Db 109063 TTTTAATATATGAACAAAATATATATTTTACTTTTGCCCTTTGTAGCATATGTCAT 109004

QY 548 ATTTCGAAGAGACAAATATATGTACA---GTTTCAAAACATCTAAGCGACAAATTA 603

Db 109003 GTTTTAAAGATTTTAAATAATTTTAAACATGTTTATGACATATATGACATTAAGAGA 108944

QY 604 TATCGAAGATATGATGACGTTGATGTTTCTTTTCATCTCTGTTATTTGTTAT 663

Db 108943 CCAAAAGTACTAATTAGCATTTTAAATGTTTATTTTATTTTATTTTATTT 108884

QY 664 GTTTTATATACATTTTCTCTCTTCAATAGATGATTTCTTCGATTTTATTAATG 723

Db 108883 AT 108824

QY 724 ACTATTAAGTCATTTTATATAGACGACGATGCTGATTTCTGTTCAAAATCTTT 783

Db 108823 TATATATTTATTAATTTATATATATATATATATATATATATATATATAT 108764

QY 784 CTGATTTTAAAGAGCTAGTTTGCAACCTGTTCTTCAAGAAATTTGATTTTTC 843

Db 108763 ATATATCATATTTAAACCTTTAATGTTAATGTTAATATATATATATATATAT 108704

QY 844 AAAAAAATTAAGTTATTTTCTTTTATTAATAAGAAACCTTGAAGAAATAGAGTTGC 903

Db 108703 AGAAAACCAAGTTAATATTTTATATATCTTATATTTTAAATTAATTAATTAATTT 108644

QY 904 CAGACTAGCCCTGATGATGTTTCCCAATTAATTAATCATACGTGATTAATTTGGC 963

Db 108643 TATTAATGCTTCATTTAGCTATTTTAAATTAATTAATTAATTAATTTTGC 108584

QY 964 AGCCCATTAATTAATTAATTAACGAACTGAATCGAGCAACCAATCGAGCTATTTC 1023

Db 108583 ACATATTAATTAATTAAGAAAGACCAAAAGTAAAAAATGTAATTTATATATCTGA 108524

QY 1024 TCTAGATTAGTAAAA 1039

Db 108523 GTTTTAAATGTTAAAA 108508

RESULT 13
PFMAL4P1 347582 bp DNA linear INV 29-JAN-2003
LOCUS Plasmodium falciparum MAL4P1.
DEFINITION AL034557 AL844503
ACCESSION AL034557.8 GI:23498126
KEYWORDS Plasmodium falciparum 3D7
SOURCE Plasmodium falciparum 3D7
ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
1
Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corcoran, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Felwell, T., Goble, A., Goodhead, I., Gilliam, R., Hamlin, N., Hane, Z., Harper, D., Hauser, H., Hornby, T., Holroyd, S., Horrocks, P., Humphrey, S., Jagels, K., James, K.D., Johnson, D., Kerhoulou, A., Knights, A., Kontoriotou, B., Kyes, S., Larke, N., Lawson, D., Leonard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moulé, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabinowitz, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sandere, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and Barrell, B.G.
Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13
JOURNAL Nature 419 (6906), 527-531 (2002)
MEDLINE 22255708

PUBMED 12368867
REFERENCE 2 (bases 1 to 347582)
AUTHORS Hamlin,N., Pain,A., Bertiman,B., Hall,N., Bowman,S., Churcher,C.,
Harris,B., Harris,D., Lawson,D., Quail,M. and Barrell,B.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On Oct 3, 2002 this sequence version replaced gi:5731897.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
source 1..347582
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P1"
10384..32958
/note="rep20 Repeats"
join(35153..41725,42757..44124)
/gene="VAR"
/note="synonym: PFD0005w"
join(35153..41725,42757..44124)
/gene="VAR"
/note="synonym: PFD0005w"
join(35153..41725,42757..44124)
/product="erythrocyte membrane protein 1 (pFEMP1)"
/protein_id="CAD49094.1"
/db_xref="GI:23498127"
/db_xref="GCA:Q81220"
/db_xref="UniProt/TREMBL:Q81220"
/translation="MVRTGGSGGVNDKGDIGHQSAKHLDSIGKKVHDQVXNGADGT
GASDAKNYIDLDKGLQKAPNINPLIGDIPCKLVEYVYNNVNGDGKRGVPTL
LSGKFPNPSTDLGGCGTNSKMRSGCEGACRYRLHCHNLESIETTSYASPTL
LLEVCMAKREGSITNHTYTKHNSKNSDPSOLCTLARSPADIGDIVAGKLPFGNT
YESAREKLENTKVEPKIHGSLSEAKKKYDGGGNYQLREDWMTANRETVMKAI
TEVKSANNYFRATGDEKNSPLTSKQCRDCKAKRP1KSGSNVAVITPTDYDFQY
LWFEEMADPCLRHKLKDAIKCKRGNGBERYCDLNRVDCNKTASQKHFVDFD
CKDCQYSCAPVDWIDNOKLEFLQAKKTYKEITSGSGSCGSGSKRRADTTTNYEVE
KKEFKELGTYKVNVNPLEKLNDEDVCTKNDIKDGNIDFNKVSAGAKKDGNNK
TFYRTKYCEACGPMWCAKVEGKWKAKENOSOTKXDYDPTKTTITELIDTTRKSDVQ
KYKPPNGCGNGEAKATPMATSRBKQKQDQMKVICYDNKKKKYSDA1NRCVL
ODGKQHRKEQKTSYNAFTFKRYVNDLHDSVERREKLSGINNAKSQCKNNKCKKE
CGCFERWVQKKEKEWAIKDFGQKQDI1EQGCDAGVTLAIVLKLEFLNDETEKSE
EKLDAEBAKEIKHLRQMLEAGVRDLAIVGPGTEGVAEQNTIMDKFLDELKSAE
OCNCKPRPAQOQPGGARSADSPAGTEDHPDAEDDDDDDDDDDEDEEEDD
POCKTNDILSTDRTKOYDCHHEKNYGNKGPWKCGDLTLVDDTKVCMPRRQKCL
VYLAHSEETKNTIEDDLRDAIRTAETPLSMQYKIKNGADAKQLDNGT1PEEFL
RSMTFTYGDYRDI CLNTDISKTYNDVAKADKIGKFPKSGSKSPGTTIPQDMQTY
GKDIWKQMI CALHGVNTKEKTKYIKNDYSYDVNQSQNGNSELDFARKPFLRMI
EMWEEFCABRGKLEQNIIGKSCNGINP1IQCSDNRHPCNACDEYKNYVETKQKEPFGQ
TTFFVADANLENADQEKDYKITQSGKQNDYLDKDCNCKSCMEGAVLIDVSSKK
PFCITVAHKSEKNCGAKFVPTNVPPAPPOPPPPALIPATPTGPNCEIVNTLP
SDNRKDACTLKYGPKAPTSWKCTPTGNTSNGAATDSGSDAKRGHNRDAPSSGS
NQSICVPPRRRLKYITPLTKAAEETTKSGKSOESKAGTSESSSEASSPGCTSSQ
GKESFGQLSTPASTSSPNSRDDLLKAVESAIVETFLMHKYKMDKLEIDDEKQK
OERSGVLGALDNGSNVDEDDKDPQKLEKSD1PEBFKQMPYTLKDIDYDIL1VRGNT
DSGANTNGSNNNNI VIEASGDQDEMKKI0KALIDEHINSLKQASAIPNORGOQOON
SSLTRETLMKEHA PSIWEQMICALTYKENDDEKIKVNDINEVEKEFGTPTGGTSGKKE
KYEYNTVYKLDENSTEAQDKTKAPASDNTPTLSHVLA PPRYRIEBSGTFCKERK
KTLAQIKVDKVDADYKCSGYSECKIEDISNTIGFADLCPGCEGCEKCRKKNM1ER
KILFEGEQSAAYVQKTKCEBSGGGNGCVTKCTQADAOFLERLGGCKANNWEG
KILFENQSBTFKTKHCDPSSFKIDCRNGKCSGTGKCDGI1T1IDAKELAKMIS
TPTVNVAVSDNDNTNTEGDDLKVCCEGKIFGKIRKEBKCRNCGGLTIDGKLGNDNG
KLDDKOIILIRALIKRMLEFLEDYKIKHKI1SDCTNNGEAGNICRDKCONKNGCEW
ITLKKEWEMKIKKHVLEKNGEGNDMKSSVRNPLKEKREHRPERNKAIYKCKGLTQPS
FCGLNGDKPSQNSHQAIDQMI KLELDTKITSCLSTSGQTEBCEHP1KCLDDEBL
LBETENPVKQOQSCFPCVKEKEKTEVDEGCKDSEPKRTDEBSKKKKDSTPSP
BPPAPATPEAPKBEKKEVPEPOPTTPOI1VDKTPALVNTSLAWSGIGFAFTYFLL

gene 45458..46601
/gene="PFD010w"
45458..46601
/note="VAR exon II, could be alternate exon II for
upstream VAR gene PFD0005w"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="synonym: PFD0015c"
complement(join(48479..49537,49768..49821))
/gene="RIF"
/note="1 probable transmembrane helix predicted by
TMHMM2.0 at aa 328-350
Similar to Plasmodium falciparum rifin pfb0030C
SMALL:O96112 (EMBL:AB001367) (370 aa) fasta scores: E(:
3.4e-65, 53.6% id in 375 aa"
/codon_start=1
/product="RIFIN"
/protein_id="CAD49095.1"
/db_xref="GI:23498128"
/db_xref="UniProt/TREMBL:Q81219"
/translation="MKIHYITLLEFLNLTILYVNRHKSNTLNTQNNRSACEELY
KAPATDDPOMKEVWVWFKSQTOQHFEYDEWVEMKQKCDKELOKILKQME
KEIMDKFATLOTIDQSDAIPTVCESKSLADKVEKGLGCGVAGVATVGLGVAV
HMYKKALEAIAKXIAIGCTADIAAALAAAGKAPMEVYIARLKGJENPFGICDT
ISSNTYNTKYATFVNTYIKTNGTGNLKRSSINPACTYITELISIKTGCACTGHP
LAIKQMTKGLAEATEAKAAEAKNKLTAIKEQYATLIEAFNSITSINASTI
ALILVILMVYIYLIRYRRKKQMKKQYIKLKE"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="synonym: PFD0020c"
complement(join(52002..53297,54200..63307))
/gene="VAR"
/note="Similar to Plasmodium falciparum variant-specific
surface protein var-3 SMALL:Q26032 (EMBL:L40609) (3006 aa)
fasta scores: E(: 6.2e-103, 43.31% id in 3551 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (pFEMP1)"
/protein_id="CAD49096.1"
/db_xref="GI:23498129"
/db_xref="GCA:Q81218"
/translation="MGTSSTPSVPKDVKNSSARVNLNIGIEIYNEKKKNGY
TSQLRGDSRRPFHDLGKARLGVIPPPANSCLDHPFTYNNNGVPPARPQDLRN
QNRRENEATONSFKIRVTGKKSAGGACAPPRONMCDKLETLIDNTDITDIDLGL
NVLYTAKYEGASIVAKHHPKENSEVCTLAASPADIGDI VSGKMFKNKEDAVQKLG
RAVFKKIDINKLEKEISIDYNDPNVYKLRPMNDVNRQVRAITCYIPYYVNFCKT
SDTFTVFNDEGCGYEGAPPTNLIDYQLEPRLMPEMBEBCRIENHLARKIKDSCRN
DKERLYGONGYDCTKRI1EKSSGSEBKCGKSCYDVQFMLEKQNOEFLOKQY
DKETIETYNKATPISNSNNTKYEKKYEELKCYQSVKVPLOLNNGRQCKETIE
EDALIFTGNKHAIRNSDYCOPCEPCVCECDGKCTCTKTDDBDKRSKTI1OKTIESE
TPIBELVY1SDKQVITETKILKDPGRGNQNNDEMLQMKVCNKNSENMKSMWLY
QDPKRYNL1LVEECGSHWAKNLI1DTIMEHQNLKNINNTVDTSCIKNCECYEA
M1ERKQDEWEKLEVLANKQDETSNHYVUKADQVRFPLQWMLADDOEKMGDFTE
DLKFKFGSVSAGTANQDA1LEFLDHLKONALTRGRNNS1KPCCTYPNPPTNGCGT
NNGNGKLYRVRIALMEMRRARKQLEKRGGINLKAADSQSKYIRGGKEKLGNOICN
IDTSS1SNSRNGNGPCTGKNDKRFK1GTMSSGEHKKKTHPEVNP1PREHFCIS
NLEKLDVSVIRKNGASHL1GDVYLLAKYBAKNIKELYQOONSNGV1IDNDKETTIC
RAMKRSFADID1IRGDMVQONTDATQYLAKEIPDKI1NHDMDIGKLCIQYNDTD
HKL1REDWENRHOVAMKCA1ENDMDCONG1P1DYI1POR1AMTBAWAVYCKE
QSLRYGELLEKCGSKGKQCTEGDVGCGKACDCKYKIDYKRNBMQMTYIKGVYKT
LYKATKTCNTTSNNPKDKQVDFYDKLLPRKSKNTGTYTAMPTNT1YSSAAGTIHQ
ELGKTVGNTKKEFQDNKRGKYAFGRHPKEYEAC1CTPRQAOQPI1KDKNDCKN1KT
LLDRNNGGTG1DGCNPK1GNVPSMNCRNSKAKENKACVAPPRKRCVSL1LAEGI
FNKKGED1RETFFVSALETYFAWKRYVNDNKKAAEELKSGT1PENFRQMYTTPADY

ROIPEGDTIDSHDIIIDVSNKANKLKEKNGEOKSVI1IDDEKLADMMKHEHIME
GMICALTHEIDEEERKAKIKSTYDOLKKTNTGTPLEKAEHPOTLRNFTESDERC
REERKEEVEVERKCKKDHEGCKNPKNGKNCVACADYSEYSTKKQNTQKEKED
IDKNGKEEYENKDKKADHYLDKCPKPGCDYMEKVNSEYWDKPKNTYTQKEDK
KCKCKPPPPAPAPACEIYVDIILNGASADYIEKCKGKYGKRPKPNCSOIH
RTNNGACMPRROKLCVYNLOYPKGTATYREAFIKCAVEFPLMHKKEPNNGSE
DLONOLSESGIIPDDPKRMGYTGDYRDLFGDIIKSGELAKKIDSLKRNIGC
KAPGDLRSKDMWNEKPYIKWMLCSIEKAWGDTIKNKSNTHYKNSDNKNGDL
ETPAKAPQFLRMFTWEDFCREOKQLODLILKCKPEKCTNGKKEKSCDACKAKE
WLOJMKENHYEKOKIKYENDDSYTNDDYQSOAYOLNKLKELIKCPGNTANBYE
KCMKYPSSQNNNMNPASLADPTSDYKDCCTCSOASRNPSVSEKSEGEPPAP
ROSLASADNPSPRAPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP
OPILARYDDE
QEEBPEPEPEVEETVAAPVPAACEIYKELPNDTKFKACULKTKGNSRLGKICI
PSGDSITTSVNDRSQNRHRAAGEATGSDAGSICVPPRRRLVGLKLTQNASQRT
OGGTSQIVKTSKASQPNSHPLPSNPBPDGADLDFI0SAAVEFLMDRYKMLTK
KPDATLGLGLOIPLAMKALNGYVPSGDNPOKKEGELPEEFKQMPETLADYDI
LFGKNDIVIGTSGASDKEMKAKEIKETIDKVPNSVSTPPPTATKPSDEKERT
MWEANGEHIMKWMIVALTYNKNGEKKIVKNEYVKKLMDANPKPKETKYOVNKLK
ENSGAKPTOPSPSGDNTPTTLNFIISRPYFYLEWGETFERKKKILBRYKSCR
GEYGEKTCGDDHDCENELKHTNMPALDORCHKQCRKRMIDIKPEEYERQK
DKYOGELKLNKNSNGNCKKEIKKHTSSEFLKELKHCKDCQNSBDDTKSEBDK
NKKIDFNKPLETFNPSYCEBPSNKNVNCNGSGRGRGKDCPTPHNEKGSWSVNA
NGNSTEITVEMIDREPLIKNSKILBESGNSDSLFKTSRLFKSVRGQWECRKYD
EKTDICKLKNPKNDKIDNOYTPFKVLEVYLQPIRGVYLKRRKIIIECKENGSTC
NENSKNDCAVCVKMAOKTTEMNOIKDHINKKEYGNGYDMSHKVKNYFEKENELKRM

Query Match 5.1%; Score 70.4; DB 3; Length 347582;
Best Local Similarity 44.6%; Pred. No. 0.0048;
Matches 275; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

QY 444 CAAAGTTTAAAAAATACCAAGTAATCACTCGAAGCTATCATGTATTGA 503
DB 335664 CAAAAGAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAATAGTAAG 335605
QY 504 AAGAACATCTATTAAAAACACGATCCTCTTAAAAAACAAGCATATTTCGAAAAGAGCA 563
DB 335604 AATATTATCTTATGATATATAAGCTATTGTGAAATTAATATTTATGAAAAATATAATAT 335545
QY 564 ATTATGTTACAGTTTACCAACATCTAAGACGCAAAATTTATCGAAAGGTAAGCTATGA 623
DB 335544 ATATATGTAATAATTTTTTTTTTATATATAATAATAATAATAATAATTTTTTTT 335485
QY 624 CGTTTCAGATTTTCTTTTCAATCTGTATATTTGTTATGTTTTATATACATTTTCTT 683
DB 335484 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 335425
QY 684 CTCTTCAATAGATGATTTCTTCGATTTTAAAAAAGCTATAAAGCTATTTTATA 743
DB 335424 TTTTTCAGTATATAAAGTATTTAACGAAATATGTAGAAAGAAAGAAATTTGGTATG 335365
QY 744 TAAAGACGACATGCTAGATTTCTCGTCAAAAATCTTTCTGATTTTTTTTAAAGCTAG 803
DB 335364 TAAATTTTATATAAATAATAAAGAAATATAATTTTTTTTAAATTAATAAATAAAG 335305
QY 804 TTTGGCAACCTGTTTCTTCAAGAAATTTGATTTTCAAAAAAATTAAGTTATTTT 863
DB 335304 TTTATTTATTTTATATAATATAATTTTTTTTAAATTTTAAATTCATATATTTTATATT 335245
QY 864 CTCTTATATAAATAGAAAACTTTAGAAAAATAGATGCGACACTAGCCCTGAAATGTT 923
DB 335244 ATATTTTTTTTTTTTTTGTCTATATATAATATAATATAATAGCTATATATAATATTTT 335185
QY 924 TTCCCAATAAATTACATCACTGTATATAATTTTGGCAGGCCCAATAATTATTTAA 983
DB 335184 TTTTTTTTTTTTTTTTAAATTTATTTTATTTTCTACACTATATAATAATAA 335125
QY 984 CCGAACTGAATCGACGAAACCAATCTGAGCTATTTCTTAGATTAGTAAAAAGGA 1043
DB 335124 TAAATTTTATTTTAAAGGTATCATTAATGTTTTTTTAAAAATAAAAAATAATAGT 335065
QY 1044 GAGAGAGAGAGAAA 1059
| | | | |

DB 335064 TATCTATATATAAAAA 335049
RESULT 14
CO422641/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 883
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 5.0%; Score 70.2; DB 6; Length 883;
Best Local Similarity 39.2%; Pred. No. 0.0065;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTTCACATTTGGCACTACCCAGAACAGATTAAAAAATACCAAGTAATCA 478
DB 873 TGTNNNNNNANCTTAAACCTTTAAANATNNNATTTAANNAAAAATATTTAAANN 814
QY 479 CTCGAAGCATCATGATATGTTTAAAGAAACATCATTAACCAACGATCCCTTAA 538
DB 813 TAAAAATTTTNNAAAACTAAATNTTAANNATNTAATTTTNTTAAAAAANNAAAA 754
QY 539 AACACGATATTTTCGAAAGAGCAAAATTTGTTACAGTTTACCAACATCTAAGAGCGACA 598
DB 753 TTAANNNTTTNTAANTATTTAAACCAAAATTTTTTTTTTAAAAAATTTTTTAAANNT 694
QY 599 AATTATATGAAAAGTAAGCTATGACGTTCAAGATTTTCTTTTCAATTTGTTATTTG 658
DB 693 AATTAATTAAATTTTNTTAATNAAAAAATTTTAAATTTTAAANCAANTTTTTT 634
QY 659 TTAATGTTTTTATATATCATTTTCTTCTCTTACATAGAGTATTTCTCGATTTTATA 718
DB 633 TTTNTTTNNATTAATAAATTTTTTAAATTAATAAANNANNTTTTTTTATNAAATATTTTA 574
QY 719 AATAGCATTAAGCATTTTATATATAAGACGACATGCTAGATTTCTCGTCAAAA 778
DB 573 AAAAAAANNTTTTTTTTTTTATANTTTAATAAATAATTTATTTNTTTTCTGNAAAAA 514
QY 779 TCTTTCTGATTTTNTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATT 838
DB 513 AAAAAAATTTTNNAAAAATNTTTTTTAAACCTTNAATNAAANNAAATTTTTTTNNNT 454
QY 839 TTTTCAAAAAAATAGATTATTTTCTCTTTAATAATGAAGAAACCTTAGAAAAATAGA 898
DB 453 NNANATTAATAAATATTTATTTTTTTTTTTTCTATTAATTTAANNANNAATTAATAAT 394
QY 899 GTTGCAGACTAGCCCTAGAAATGTTTTCCCAATAATTAATCACTGTATTAATTT 958
DB 393 TTTNAAAAAATAATTTTATANTTTATNTNANAAAAATTTTAAANTNAAANNAAAAATTTTA 334
QY 959 TGGCCAGCCCAATAATTTTAAACGAAATCTGAAATCGAGGAAACCAATCTGAGCT 1018
DB 333 NTTTAATTAATTAATAATTAATTTTAAACATCTTTTAAATTTTATATTAATTAATAAATAAATAAT 274
QY 1019 ATTT 1022
| | |

Db 273 NTT 270

RESULT 15
LOCUS BV119878/c 241 bp DNA linear STS 19-MAR-2004
DEFINITION PZA01377 CML247 Zea mays CML247 Zea mays STS genomic, sequence
cagged site.
ACCESSION BV119878
VERSION BV119878.1 GI:45589251
KEYWORDS STS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 241)
McMullen,M.D., Vroch Bi.,I., Schroeder,S.S. and Gaut,B.S.
MPZ-UCI Joint SNP Discovery
Unpublished (2003)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Contact: Brandon S. Gaut
Dept. Ecology and Evolutionary Biology
U.C. Irvine
321 Steinhilber Hall, Irvine, CA 92697-2525, USA
Tel: (949) 824-2564
Fax: (949) 824-2181
Email: bgaut@uci.edu
Primer A: TTCTACGTCGAATGATCATCAAGA
Primer B: AGGAACCAAGCCTTATGAGAGA
STS size: 241
Protocol:
PCR amplification of genomic DNA
Template: 50 ng
Primer: each 0.5 uM
dNTPs: each 200 uM
Taq Polymerase: RedTaq (Sigma)
Total Vol: 10 uL
Amplicon sequencing
ABI protocol - using d-Rhodamine terminator cycle
Sequencing ready reaction with ampliTaq DNA polymerase FS
Sequencing ran on ABI 3700 sequencer.

Buffer:
Genomic DNA amplification
RedTaq (Sigma)
Sequencing buffer
d-Rhodamine Kit (ABI)

PHRED/PRAP Quality Scores 56 56 56 51 51 40 40 40 40 35 35 42 48
56 56 56 56 42 42 56 33 33 33 47 44 37 37 37 37 37 37 48 48
42 42 42 42 30 30 30 37 37 44 44 44 84 90 89 81 71 68 51 40 45
45 51 39 39 39 56 44 48 86 86 90 81 71 68 63 62 62 58 58 69 69
67 67 77 72 68 71 68 72 74 80 74 76 73 66 77 79 68 62 57 56 53 60
58 71 71 68 68 77 75 79 71 79 68 81 75 90 79 35 37 33 39 42 42 35
37 30 38 20 48 40 44 40 44 33 42 37 74 68 60 50 54 54 51 51 55
49 50 57 57 72 82 90 90 88 88 79 74 74 67 64 62 53 50 52 59 55
65 71 71 84 84 84 88 85 90 84 88 88 90 80 79 83 83 86 86 81
76 68 49 48 39 39 42 49 33 42 29 24 24 14 14 14 32 44 40 48 48 90
84 84 84 87 79 72 72 63 57 53 51 41 57 54 56 56 49 51 30 33 33 28
28 28 33 40 42 48 37 37.

FEATURES

source
1. .241
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="CML247"
/db_xref="taxon:4577"
/clone_lib="Zea mays CML247"
/dev_stage="seedling"
/note="Organ: leaf; genomic DNA from inbred line"
<1. .>241

STS
ORIGIN

Query Match 5.0%; Score 70; DB 11; Length 241;
Best Local Similarity 61.5%; Pred. No. 0.00074;
Matches 112; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
Qy 799 GCTAGTTGGCAACCCGTTTCTTCAAGAAATTTTGAATTTTCAAAAAAATTAGTTT 858
Db 195 GTTAGTTGGAAACCAATTTTCCCAAGGAAATTTCAATTTCCCAAGGAAATTAAGTTC 136
Qy 855 ATTTTCTCTTTATTAATTAAGAAACATTAAGAAATTAAGCTTGCACACTAGCCCTAGA 918
Db 135 ATTTTCCCTTGGAAATTAAGAAATCCCAAGGAAATGAGATTCCCAACTTAACCTTAAT 76
Qy 919 ATGTTTCCCAATTAATTACATCACTGTGTATTAATTTTGGCCAGCCCAATTAATAT 978
Db 75 ACTGTTAATTTAGTTCCCAATATCCATCACCAAGATGATATCTTAACCGCATATCTGT 16
Qy 979 TT 980
Db 15 TT 14

Search completed: September 16, 2005, 03:01:11
Job time : 4684.65 secs

XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ

Query March	100.0%	Score 1394	DB 2	Length 1394
Best Local Similarity	100.0%	Pred. No. 5.1e-291		
Matches 1394	Conservative	0	Mismatches 0	Indels 0
				Gaps 0

Qy	1	CCAGGAGTCTCTTATGAAAAAGATAGTATACATAGTGTCTATATCCGTTTTCTTTAGGGTCC	60
Db	1	CCATGGTGTCTCTATGAAAAAGATAGTATCAATGTGTCTATATCCGTTTTCTTTAGGGTCC	60
Qy	61	CTTCTTCTGCTTATTTACTGACTGAATCGGGGTATCAAAAACTTTCACGGGTGCATGAT	120
Db	61	CTTCTTCTGCTTATTTACTGACTGAATCGGGGTATCAAAAACTTTCACGGGTGCATGAT	120
Qy	121	CTCATGTGTCACTTCTCCCACTCCGCGTGTGCAATTTCTTGGATGTGGTGGTCCCAT	180
Db	121	CTCATGTGTCACTTCTCCCACTCCGCGTGTGCAATTTCTTGGATGTGGTGGTTCCAT	180
Qy	181	CTGACCGAGGCCCATCAGACACCTTTCCGGGACACCCATCAAGGGCTTTCCGATGGCCCA	240
Db	181	CTGACCGAGGCCCATCAGACACCTTTCCGGGACACCCATCAAGGGCTTTCCGATGGCCCA	240
Qy	241	CGAAGCGTATCGGGTCTGGTGTATCCAGGGGATATATGTCCGCCCAATTCGTCACTTAT	300
Db	241	CGAAGCGTATCGGGTCTGGTGTATCCAGGGGATATATGTCCGCCCAATTCGTCACTTAT	300
Qy	301	TTATTTATCTTTAGATATATTTTAAATTTTGGAAAAATTAACAACTTTATCTTTGTGTAT	360
Db	301	TTATTTATCTTTAGATATATTTTAAATTTTGGAAAAATTAACAACTTTATCTTTGTGTAT	360
Qy	361	GGGCTTCAGCAATGATTTTCGCTTAAAGGCCCAAAATCGAGGACCAAGCATGTCTAGTG	420
Db	361	GGGCTTCAGCAATGATTTTCGCTTAAAGGCCCAAAATCGAGGACCAAGCATGTCTAGTG	420
Qy	421	TCCACTATTTGGCATCACCCAGAACAAAGTTTAAAAATTAACCAAGTATCTAATCCACT	480
Db	421	TCCACTATTTGGCATCACCCAGAACAAAGTTTAAAAATTAACCAAGTATCTAATCCACT	480
Qy	481	CGAAAGCTATCATGTATATGTTTAAAGAAACATCTATTAAAAACACAGATCCTTTAAAAAA	540
Db	481	CGAAAGCTATCATGTATATGTTTAAAGAAACATCTATTAAAAACACAGATCCTTTAAAAAA	540
Qy	541	CAAGCATTTTCGAAAGAGACAAATATATGTTACAGTTTACAAACATCTAAGAGCACAAC	600
Db	541	CAAGCATTTTCGAAAGAGACAAATATGTTTACAGTTTACAAACATCTAAGAGCACAAC	600
Qy	601	TTATATCGAAAGTATGACTATGACGTTAGATTTTCTTTTCATCTCTGTATATTTGGT	660
Db	601	TTATATCGAAAGTATGACTATGACGTTAGATTTTCTTTTCATCTCTGTATATTTGGT	660
Qy	661	ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTTCGATTTATATA	720
Db	661	ATTGTTTTTATATACATTTTCTTCTCTTACAAATAGATGATTTTCTTCGATTTATATA	720
Qy	721	ATGACTATAAGTATTTTATATATPAGAGCAGCATGTCGTATGTCGGTCAAAAAATC	780
Db	721	ATGACTATAAGTATTTTATATATPAGAGCAGCATGTCGTATGTCGGTCAAAAAATC	780
Qy	781	TTTTCTGATTTTTTAAAGCTAGTTGGCAACCCGTGTTCTTTCMAAGATTTTGATTTT	840
Db	781	TTTTCTGATTTTTTAAAGCTAGTTGGCAACCCGTGTTCTTTCMAAGATTTTGATTTT	840
Qy	841	TTTCAAAAAAATATGTTTATTTTCTCTTTATTAATAAGAAAAACATTGAAAAAATAGGT	900
Db	841	TTTCAAAAAAATATGTTTATTTTCTCTTTATTAATAAGAAAAACATTGAAAAAATAGGT	900
Qy	901	TGCGACAGCTAGCCCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTGTATATATTTGG	960
Db	901	TGCGACAGCTAGCCCTAGAAATGTTTTCCCAATTAATTAACAATCACTGTGTATATATTTGG	960
Qy	961	GCCAGCCCCATTAATATTTTAAACCGAAACTGAAATCGAGGAAACCAATCTGAGCTAT	1020
Db	961	GCCAGCCCCATTAATATTTTAAACCGAAACTGAAATCGAGGAAACCAATCTGAGCTAT	1020

Db	961	GCCAGCCCCATTAAATTATTTTAAACCBAATCGAAATCGACGAAACCAAAATCTGAGCTAT	1020
Qy	1021	TTCTCTAGATTATGTAATAAGAGAGAGAGAGAGAAATCAAGTTTAACTGATTTGCTGCC	1080
Db	1021	TTCTCTGATTAGTAAATAAGGAGAGAGAGAGAAAGAAATCAGTTTAACTGATTTGCTGCC	1080
Qy	1081	TGAGATGTGGGGTTTGGCAACGATAGCACCGTAATATATAGCTCATAGAGTGCCATGCA	1140
Db	1081	TGAGATGTGGGGTTTGGCAACGATAGCACCGTAATATATAGCTCATAGAGTGCCATGCA	1140
Qy	1141	GGTTCCGACAGCTCTCGTGTATCTCAACATGGAGCAATACTACATGCTTGTTCAACCGTTGCTC	1200
Db	1141	GGTTCCGACAGCTCTCGTGTATCTCAACATGGAGCAATACTACATGCTTGTTCAACCGTTGCTC	1200
Qy	1201	TTGTTTCATGTCGTCAAGCCTTGGCTATTTCTGAACCAAGAGATACCTACTGCCAACAAT	1260
Db	1201	TTGTTTCATGTCGTCAAGCCTTGGCTATTTCTGAACCAAGAGATACCTACTGCCAACAAT	1260
Qy	1261	CCATCTTACCTCATGCAACTTCATGCAAAACAGCACATATGTTCCGAAACCAATGCATT	1320
Db	1261	CCATCTTACCTCATGCAACTTCATGCAAAACAGCACATATGTTCCGAAACCAATGCATT	1320
Qy	1321	AAAGATCAACAACAGCTAAGCGTTCTCCGCTAGAGCTTCCCTCTCTCTCTGCGAGATCTTTTT	1380
Db	1321	AAAGATCAACAACAGCTAAGCGTTCTCCGCTAGAGCTTCCCTCTCTCTCTGCGAGATCTTTTT	1380
Qy	1381	CGTCCACACCACTATG 1394	
Db	1381	CGTCCACACCACTATG 1394	

RESULT 2
AAH76332
ID AAH76332 standard; DNA; 1394 BP

AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
PN
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albersen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

CC Ms45), which encodes a product selected from auxins, rolB and diptheria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;

Query Match 99.7%; Score 1389.2; DB 2; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 5,6e-290;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 CCAATGCTCTCTATGAAAAAGATGAGTACATGTCATATCCGTTTCTTAAGGCTC 60
DB 1 CCAATGCTCTCTATGAAAAAGATGAGTACATGTCATATCCGTTTCTTAAGGCTC 60
QY 61 CTTCTTCTGCTTATTAATGATCGGGGTTACAAAAACTTCCACGGGTGCATGAT 120
DB 61 CTTCTTCTGCTTATTAATGATCGGGGTTACAAAAACTTCCACGGGTGCATGAT 120
QY 121 CTCATGTTTCCACTTCTCCACCTCGGTTGCACATTTCTGGATGTCGGTCCCAT 180
DB 121 CTCATGTTTCCACTTCTCCACCTCGGTTGCACATTTCTGGATGTCGGTCCCAT 180
QY 181 CTGACCGAGGCCCATCATGACACCTTTCGGGACACCCATCAAGGGCTTTGGAATGACCA 240
DB 181 CTGACCGAGGCCCATCATGACACCTTTCGGGACACCCATCAAGGGCTTTGGAATGACCA 240
QY 241 CGAGAGTATCGGGTGTGGTATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
DB 241 CGAGAGTATCGGGTGTGGTATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
QY 301 TTATTAATCTTTAGATATTAATTAATTTTGGAAAAATACAATTAATCTTTTGTGA 360
DB 301 TTATTAATCTTTAGATATTAATTAATTTTGGAAAAATACAATTAATCTTTTGTGA 360
QY 361 GGGCCTTCAGATATGATTTTGGCTTAAAGGCCCAAAATCGAGAGACCAACCATCTAATG 420
DB 361 GGGCCTTCAGATATGATTTTGGCTTAAAGGCCCAAAATCGAGAGACCAACCATCTAATG 420
QY 421 TCCACTATTGGGACTCCCGAACAAGATTAAAAAAATAACAAAGTAACTAATCCACT 480
DB 421 TCCACTATTGGGACTCCCGAACAAGATTAAAAAAATAACAAAGTAACTAATCCACT 480
QY 481 CGAAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACCAGATCCTCTTAAAAA 540
DB 481 CGAAAGCTATCATGTAATGTTTAAAGAACATCTATTTAAACCAGATCCTCTTAAAAA 540
QY 541 CAAGCATATTTGCAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAAGCGACAAA 600
DB 541 CAAGCATATTTGCAAGAGACAAATTAATGTTTACAGTTTACAAACATCTAAAGCGACAAA 600
QY 601 TTATATCGAAAGGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTTGTTATTTGTT 660
DB 601 TTATATCGAAAGGTAAGCTATGACGTTACAGATTTTCTTTTCAATCTTGTTATTTGTT 660
QY 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGAATTTTCTTCGATTTTAA 720
DB 661 ATTGTTTTTATATACATTTTCTTCTTACATAGATGAATTTTCTTCGATTTTAA 720
QY 721 ATGACTATTAAGTCAATTTTATATAGAGACGCAATGCTGTAATCTGTTCAAAAATC 780
DB 721 ATGACTATTAAGTCAATTTTATATAGAGACGCAATGCTGTAATCTGTTCAAAAATC 780
QY 781 TTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
DB 781 TTCTGATTTTTTAAAGAGCTAGTTGGCAACCTGTTTCTTCAAGAATTTGATTTT 840
QY 841 TTCAAAAAAATTAAGTTATTTCTCTTTAATAAATAGAAAAACATTAGAAAAATGAGT 900
DB 841 TTCAAAAAAATTAAGTTATTTCTCTTTAATAAATAGAAAAACATTAGAAAAATGAGT 900
QY 901 TGGCAGACTAGCCCTAGAAATGTTTTGCCAATAAATTAATCACTGCTGTAATTAATTTTG 960
DB 901 TGGCAGACTAGCCCTAGAAATGTTTTGCCAATAAATTAATCACTGCTGTAATTAATTTTG 960

```

```

QY 961 GCCAGCCCATTAATTTTAAACGAAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
DB 961 GCCAGCCCATTAATTTTAAACGAAACTGAAATCGAGCGAAACCAATCTGAGCTAT 1020
QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATGTCCTCC 1080
DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAAATCAGTTTAAAGTATGTCCTCC 1080
QY 1081 TGAATGTGCGGTTTGGCAACATAGCCACCGTAATCATAGCTCATAGTGCTTACGTCA 1140
DB 1081 TGAATGTGCGGTTTGGCAACATAGCCACCGTAATCATAGCTCATAGTGCTTACGTCA 1140
QY 1141 GGTGGGAGCTCTGCTGTCATCTGACATGAGATCTCATAGCTTGTTCACGTTCCGC 1200
DB 1141 GGTGGGAGCTCTGCTGTCATCTGACATGAGATCTCATAGCTTGTTCACGTTCCGC 1200
QY 1201 TTGTTTCATTCGTCAAGCCTTGCCTATTCTGAACCAAGAGATTAACCTCCAAACAT 1260
DB 1201 TTGTTTCATTCGTCAAGCCTTGCCTATTCTGAACCAAGAGATTAACCTCCAAACAT 1260
QY 1261 CCATCTTACTCATGCACTTTCATGCAACACGCAATATGTTTCTGAAACCAATTCAT 1320
DB 1261 CCATCTTACTCATGCACTTTCATGCAACACGCAATATGTTTCTGAAACCAATTCAT 1320
QY 1321 AAAAGATCAACAAGCTAGGCTTCCCGCTAGCTTCCGCTCTGCTGCGATCTTTT 1380
DB 1321 AAAAGATCAACAAGCTAGGCTTCCCGCTAGCTTCCGCTCTGCTGCGATCTTTT 1380
QY 1381 CGTCCACCAACCATG 1394
DB 1381 CGTCCACCAACCATG 1394

```

RESULT 4
 AAH76333
 ID AAH76333 standard; DNA; 1394 BP.
 AC AAH76333;
 DT 29-OCT-2001 (first entry)
 XX
 DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
 XX
 KW Ms45; male tissue; regulatory region; transcription; male fertility;
 KW hybrid seed; ds.
 XX
 OS Zea mays.
 XX
 PN WO200160997-A2.
 PD 23-AUG-2001.
 XX
 PF 13-FEB-2001; 2001WO-US004527.
 XX
 PR 15-FEB-2000; 2000US-00504487.
 XX
 PA (PION-) PIONEER HI-BRED INT INC.
 PI Albersen MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;
 XX
 DR WPI; 2001-514772/56.
 XX
 PT A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the MS45 gene useful for
 PT mediating fertility in a male plant.
 XX
 PS Claim 4; Page 47; Sopp; English.
 XX
 CC The invention provides a male tissue-preferred regulatory region (1)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the MS45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a

CC the MS45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment

XX
SO Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 17.1%; Score 238.8; DB 5; Length 255;
Best Local Similarity 97.2%; Pred. No. 7.5e-42;
Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1145 CGGAGCTCGTGTGATCATTCACATGGCATACATACATGCTTTCAACCGTGTCTTGG 1204

DB 5 CCGGGATTCGCGGTGATCATTCACATGGCATACATACATGCTTTCAACCGTGTCTTGG 64

QY 1205 TCCATCGTCCAGCTTCCCTTATTTGGAACCAAGAGATACCTACTCCCAACATTCAT 1264

DB 65 TCCATCGTCCAGCTTCCCTTATTTGGAACCAAGAGATACCTACTCCCAACATTCAT 124

QY 1265 CTTACTCATGCACTTTCATGCAACACGCAATATTTTCTGAAACCAATTCATTAAAG 1324

DB 125 CTTACTCATGCACTTTCATGCAACACGCAATATTTTCTGAAACAGATCTATTAAAG 184

QY 1325 ATCACAACAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTGCGCGATCTTTTCGTC 1384

DB 185 ATCACAACAGCTAGCGCTTCCCGCTAGCTTCCCTCTCTGCGCGATCTTTTCGTC 244

QY 1385 CACCACCATTG 1394

DB 245 CACCACCATTG 254

RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.

XX
AC AAH76334;

DT 29-OCT-2001 (first entry)

XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.

XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;

XX
KM hybrid seed; ds.

XX
OS Zea mays.

XX
PN WO200160997-A2.

XX
PD 23-AUG-2001.

PF 13-FEB-2001; 2001WO-US004527.

XX
PR 15-FEB-2000; 2000US-00504487.

XX
PA (PION-) PIONEER HI-BRED INT INC.

XX
PI Albertsen MC, Fox TW, Garmaat CW, Huffman G, Kendall TL;

XX
DR WPI; 2001-514772/56.

PT A male tissue-preferred regulatory region comprising nucleotide sequences

PT essential for initiating transcription of the Ms45 gene useful for

XX
PS Claim 5; Page 47; 50pp; English.

XX
CC The invention provides a male tissue-preferred regulatory region (I)

CC comprising nucleotide sequences essential for initiating transcription of

CC the Ms45 gene. A method of mediating male fertility in a plant is

CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence

XX
SO Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;

Query Match 10.5%; Score 146; DB 5; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.1e-22;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CCGTGCATTCACATGGCATTACTAGTGTTCGAAACCGTTCGTC-TTGTCCATTCGC 1213

DB 1 CCGTGCATTCACATGGCATTACTAGTGTTCGAAACCGTTCGTC-TTGTCCATTCGC 60

QY 1214 CAAGCTTGCCTATTTCTGAACCAAGAGATACCTACTCCCAACATTCATCTACTCAT 1273

DB 61 CAAGCTTGCCTATTTCTGAACCAAGAGATACCTACTCCCAACATTCATCTACTCAT 120

QY 1274 GCAACTTCCATGCAACACGCAATATGTTTCTGAAAC 1311

DB 121 GCAACTTCCATGCAACACGCAATATGTTTCTGAAAC 158

RESULT 7
AAL15210/C

ID AAL15210 standard; cDNA; 883 BP.

XX
AC AAL15210;

DT 07-DEC-2001 (first entry)

XX
DE Human breast cancer expressed polynucleotide 7667.

XX
KM Human; breast cancer; cell marker; cytostatic; ss.

XX
OS Homo sapiens.

XX
PN WO200151628-A2.

XX
PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US0000798.

XX
PR 14-JAN-2000; 2000US-0176077P.

XX
PR 14-MAR-2000; 2000US-0189167P.

XX
PR 24-MAR-2000; 2000US-0192093P.

XX
PR 29-MAR-2000; 2000US-0193480P.

XX
PR 15-MAY-2000; 2000US-0205230P.

XX
PR 09-JUN-2000; 2000US-0211315P.

XX
PR 25-JUL-2000; 2000US-0220534P.

XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX
DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.

XX
PS Claim 1; Page 1378; 3695pp; English.

XX
CC The invention relates to human breast cancer expressed polynucleotides

CC (AAL07544-AAL26789) and methods of assessing whether a patient is

XX Sequence 883 BP; 322 A; 18 C; 23 G; 396 T; 0 U; 124 Other;
SQ
Query Match 5.0%; Score 70.2; DB 4; Length 883;
Best Local Similarity 39.2%; Pred. No. 2.6e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTGGCAGTACCCAGAACAGATTAAAAAAATTAACCAAGTAATCCCA 478
DB 873 TGTNNNNNNNACCTNACCTTTAAANATNNNANTNNANNNAAAAATATTTAANN 814
QY 479 CTCGAAGCTATCATGTATGTTTAAAGAAACATCTATAAACCGATCCTTAAAA 538
DB 813 TAAAAATTTTTTNNAAAACTAATTNTAANNTAATNTTNTAAAAAANNAAAA 754
QY 539 AACAGCATATTTGGAAGAGACAAATTATGTACAGTTTACAAACATCTAGAGCGACA 598
DB 753 TTAANNNTTNTAANTTATTAACCAAAATTTTTTTAAAAAATTTTTTAAANNT 694
QY 599 AATTATATCGAAAGTATGAGTACGTTGAGATTTTCTTTTCATTTCTGTATTG 658
DB 693 AATTAATTTNAATTTTTTTTAAATNAAAAAATTTTAAATTTTAAACAAANTTTTTT 634
QY 659 TTAATGTTTTTATATACATTTTCTCTCTACATATAGATTTTCTCGATTTTATA 718
DB 633 TTTNTTTNNATTAATAAATTTTAAATTAATAAANNANNTTTTTTTATNAAATATTTA 574
QY 719 AAATGACTATMAAGCTATTTTATATATAGACAGCATGTCTGATTCCTGTTCAAAA 778
DB 573 AAAAAAANNNTTTTTTTTATANTTTATATAAATAATTTATNTNTTTTTCTNAAAA 514
QY 779 TCTTTCTGATTTTTTTAAGCTAGTTGGCAACCCGTCTTTCTTCAAGATTTTGATT 838
DB 513 AAAAAAATTTTTTNNAAATNTTTTTTAAACCTTNAATTAANANANATTTTTTNNNT 454
QY 839 TTTTCAAAAAATTTAGTTATTTTCTCTTATATAATAGAAAACTTGAATAATAGA 898
DB 453 NNNANNTTAAAAATTTATTTTNTTANTTTATTAATTAATAAANANANATTAAT 394
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGTATATAT 958
DB 393 TTTNAAAAAATAATTTATANTTATNTAANANAAAAATTTAAANTNANAAAAATTTTA 334
QY 959 TGGCCAGCCCATTAATTTATTAACCGAAATCGAAATCGAGGAAACCAATCTGAGCT 1018
DB 333 NTTTAATTAATAAATAATTAATTAACNTCATTTNTTATTTTATTAATTAATAAAAAAAT 274
QY 1019 ATTT 1022
DB 273 NTTT 270

RESULT 8
ACN85231/c
ID ACN85231 standard; DNA; 960 BP.
AC ACN85231;
XX 02-DEC-2004 (first entry)
XX
DE Breast cancer related marker, seq id 6381.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; dr.
XX
XX Homo sapiens.
XX OS
XX PN US2003099974-A1.
XX
XX PD 29-MAY-2003.
XX
XX PF 18-JUL-2002; 2002US-00198846.
XX
XX PR 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.
PA
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI
XX WPI; 2003-787014/74.
DR
XX
PT Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 6381; 36pp; English.
PS
XX The invention relates to an isolated polypeptide (I) associated with
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?docID=2003099974
XX
SQ Sequence 960 BP; 340 A; 33 C; 39 G; 421 T; 0 U; 127 Other;
Query Match 5.0%; Score 70.2; DB 11; Length 960;
Best Local Similarity 39.2%; Pred. No. 2.6e-05;
Matches 237; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 419 TGTCCACTATTGGCAGTACCCAGAACAGATTAAAAAAATTAACCAAGTAATCCCA 478
DB 933 TGTNNNNNNNACCTNACCTTTAAANATNTNNNANTNNANNNAAAAATATTTAANN 874
QY 479 CTCGAAGCTATCATGTATGTTTAAGAAACATCTATTAACCGATCCTTAAAA 538
DB 873 TAAAAATTTTTTNNAAAACTAATTNTAANNTAATNTTNTAAAAAANNAAAA 814
QY 539 AACAGCATATTTGGAAGAGACAAATTATGTACAGTTTACAAACATCTAGAGCGACA 598
DB 813 TTAANNNTTNTAANTTATTAACCAAAATTTTTTTTAAAAAATTTTTTAAANNT 754
QY 599 AATTATATCGAAAGTATGAGTACGTTGAGATTTTCTTTTCATTTCTGTATTG 658
DB 753 AATTAATTTNAATTTTTTTTAAATNAAAAAATTTTAAATTTTAAACAAANTTTTTT 694
QY 659 TTAATGTTTTTATATACATTTTCTCTCTACATATAGATTTTCTCGATTTTATA 718
DB 693 TTTNTTTNNATTAATAAATTTTAAATTAATAAANNANNTTTTTTTATNAAANATTTTA 634
QY 719 AAATGACTATMAAGCTATTTTATATATAGACAGCATGTCTGATTCCTGTTCAAAA 778
DB 633 AAAAAAANNNTTTTTTTTATANTTTATATAAATAATTTATNTNTTTTTCTNAAAA 574
QY 779 TCTTTCTGATTTTTTTAAGCTAGTTGGCAACCCGTCTTTCTTCAAGATTTTGATT 838
DB 573 AAAAAAATTTTTTNNAAATNTTTTTTAAACCTTNAATTAANANANATTTTTTNNNT 514
QY 839 TTTTCAAAAAATTTAGTTATTTTCTCTTATATAATAGAAAACTTGAATAATAGA 898
DB 513 NNNANNTTAAAAATTTATTTTNTTANTTTATTAATTAATAAANANANATTAATTAAT 454
QY 899 GTTGCCAGACTAGCCCTAGAAATGTTTCCCAATTAATTAATCACTGTGTATATAT 958
DB 453 TTTNAAAAAATAATTTATANTTATNTAANANAAAAATTTAAANTNANAAAAATTTTA 394
QY 959 TGGCCAGCCCATTAATTTATTAACCGAAATCGAAATCGAGGAAACCAATCTGAGCT 1018
DB 393 NTTTAATTAATAAATAATTAATTAACNTCATTTNTTATTTTATTAATTAATAAAAAAAT 334
QY 1019 ATTT 1022

Db 333 NTTT 330

RESULT 9
ADRo4296
ID ADRo4296 standard; DNA; 13400 BP.

XX ADRo4296;

DT 04-NOV-2004 (first entry)

DE Corn FT homologue nucleotide sequence SEQ ID NO:63.

XX flowering locus T gene; FT; terminal flower; TFL; Ap3; plant;
KM floral development; plant sterility; plant fertility; flowering time;
KM plant growth rate; inflorescence architecture; tissue culture morphology;
KW cell division; FT homologue; gene; ds.

XX Zea mays.

OS WO2004067723-A2.

PN 12-AUG-2004.

XX 29-JAN-2004; 2004WO-US002422.

PF 30-JUN-2003; 2003US-00343477.

PR (PION-) PIONEER HI-BRED INT INC.
PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX Danilevskaya O, Hermon P, Bruggemann E, Shibrour D, Ananiev E;
PI Rafalski JA, Sakai H, Cahoon E, Cahoon R, Klein T;

XX WPI; 2004-580996/56.

PT New polynucleotides, specifically nucleic acid fragments encoding
PT flowering locus T gene (FT) or terminal flower (TFL), or Apetala3 (Ap3)
PT homologue, useful for floral development, e.g. engineering plant flowering
PT time.

PS Claim 6; SEQ ID NO 63; 109pp; English.

XX The present invention describes an isolated polynucleotide comprising a
CC first, second, third, fourth or fifth nucleotide sequence, or their
CC complement encoding a polypeptide either having flowering locus T gene
CC (FT), terminal flower (TFL), or Apetala3 (Ap3) homologue activity. Also
CC described: (1) a vector comprising the polynucleotide; (2) a recombinant
CC DNA construct comprising the polynucleotide; (3) transforming a cell by
CC transforming a cell with the polynucleotide; (4) a cell comprising the
CC recombinant DNA construct; (5) producing a plant comprising transforming
CC a plant cell with the polynucleotide, and regenerating a plant from the
CC transformed plant cell; (6) a plant comprising the recombinant DNA
CC construct; (7) a seed comprising the recombinant DNA construct; (8) an
CC isolated polynucleotide comprising a first nucleotide sequence, where the
CC first nucleotide sequence contains at least 30 nucleotides, and where the
CC first nucleotide sequence is comprised by another polynucleotide, where
CC the other polynucleotide includes the second, third, fourth, fifth or
CC sixth nucleotide sequence; (9) an isolated polypeptide having FT or Ap3
CC homologue activity, as described above; and (10) isolating a polypeptide
CC encoded by the polynucleotide comprising isolating the polypeptide from a
CC cell containing a recombinant DNA construct comprising the polynucleotide
CC operably linked to a regulatory sequence. The polynucleotides are useful
CC for floral development, e.g. engineering plant sterility/fertility,
CC flowering time, plant growth rate, inflorescence architecture, and tissue
CC culture morphology and the rate of cell division to enhance
CC transformation. The present sequence represents an FT homologue
CC nucleotide sequence from the present invention.

XX Sequence 13400 BP; 3962 A; 2948 C; 2695 G; 3795 T; 0 U; 0 Other;

Query Match 5.0%; Score 69.8; DB 13; Length 13400;

Best Local Similarity 73.6%; Pred. No. 6.3e-05;
Matches 89; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 811 ACCCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTA 870

Db 8559 ACCATTTTATTTTCAAGAGTTTATTTATCAAGAAAAATAGTCTCTCTTG 8618

QY 871 TAAATAGAAAACCTTGAAAAATAGAGTCCGACGCTAGAGATGTTTCCCA 930

Db 8619 AAAAATAAAAATCATTAGAAAAATGGGCTGTCAAACTGCTTATTAGTTTCAT 8678

QY 931 T 931

Db 8679 T 8679

RESULT 10
AA210551
ID AA210551 standard; DNA; 2657 BP.

XX AA210551;

DT 16-NOV-1999 (first entry)

DE DNA sequence of the P-Zea promoter of maize.

XX P-Zea promoter; male sterile plant; glyphosate tolerance; glyphosate;
KM male reproductive tissue; hybrid seed production; crop outcrossing;
KW flower life; ss.

OS Zea mays.

PN WO9946396-A2.

PD 16-SEP-1999.

XX 09-MAR-1999; 99WO-US005126.

PR 09-MAR-1998; 98US-0077277P.

XX (MONS) MONSANTO CO.

PI Brown SM, Fromm ME;

DR WPI; 1999-551420/46.

PT Production of male sterile plants using a gene encoding glyphosate
PT tolerance, used for, e.g. production of hybrid seed.

PS Disclosure; Fig 1A-B; 54pp; English.

XX The present sequence represents the P-Zea promoter of maize. It is used
CC in the method of the invention. The specification describes a method for
CC the production of male sterile plants. The method comprises selective
CC expression of DNA encoding a protein that causes tolerance to glyphosate
CC and application of glyphosate. The method uses two DNA molecules, each
CC operably linked to a separate promoter, whereby the first promoter
CC functions in plant cells to produce a protein that causes tolerance to
CC glyphosate, and the second promoter functions in plant cells to cause the
CC production of a second RNA sequence in a male reproductive tissue.
CC Expression of the DNA promotes tolerance to glyphosate in those tissues
CC in which it is expressed. Expression of the second DNA molecule causes
CC the production of an RNA sequence which can inhibit the glyphosate
CC tolerance generated by expression of the first DNA molecule. By using a
CC promoter for the second DNA molecule which restricts the production of
CC the antisense RNA to only a subset of the tissues which express the first
CC DNA molecule, only the subset of tissues in which the second DNA molecule
CC is expressed will be susceptible to glyphosate toxicity. In this way, a
CC specific cell type or combination of cell types, depending upon the
CC promoters utilized, can be selectively ablated by application of
CC glyphosate to the plant. The methods can be used for producing male-
CC sterile plants for use in the production of hybrid seed, for minimizing
CC undesirable crop outcrossing, and for lengthening flower life. The

CC methode can be used with plants such as corn, wheat, rice, canola, oat,
CC barley, alfalfa, carrot, cotton, oilseed, oilseed rape, sugarbeet,
CC sunflower, soybean, tomato, cucumber and squash
XX

Sequence 2657 BP; 686 A; 611 C; 677 G; 683 T; 0 U; 0 Other;

Query Match 4.8%; Score 67.2; DB 2; Length 2657;

Best Local Similarity 74.2%; Pred. No. 0.00015; Mismatches 33; Indels 1; Gaps 1;

Matches 98; Conservative 0; Mismatches 33; Indels 1; Gaps 1;

DB 1883 ATCTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGATTGTTTTCAA 846

DB 1883 ATCTTTTAAAGCTAGTTGGCAACCTGTTCTTTCAAGATTGTTTTCAA 1941

DB 1942 GGAATATAGTCAATTTCCCTTGGAATAATATGGAATCCATGGAATAATGTCCTCA 2001

DB 907 ACTAGCCCTAGA 918

DB 2002 ACTAGCCCTAAA 2013

RESULT 11

AAKS8751/c

ID AAKS8751 standard; cDNA; 6027 BP.

AC AAKS8751;

DT 16-AUG-1999 (first entry)

DE Maize d111 gene encoding starch synthase enzyme D11.

XX Starch synthase; SSI; D11; d111 gene; maize; transgenic plant; ss.

OS Zea mays.

XX Key

FT m18c_feature

FT 1.1137

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 16"

FT 120.5147

FT /tag= a

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 12"

FT 565.816

FT /tag= d

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 14"

FT 655.1221

FT /tag= c

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 13"

FT 1369.1944

FT /tag= e

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 15"

FT 1438.2424

FT /tag= g

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 17"

FT 2425.3791

FT /tag= h

FT /note= "Functional fragment of cDNA specifically claimed

FT in Claim 18"

XX WO924575-A1.

XX PN 20-MAY-1999.

XX PD

XX XX

PF 12-NOV-1998; 98WO-US024225.

XX 12-NOV-1997; 97US-00968542.

XX (IOWA) UNIV IOWA STATE RES FOUND INC.

XX Myers AM, James MG;

XX WPI; 1999-327406/27.

XX P-PSDB; AAY06199.

XX Nucleic acid encoding starch synthase enzyme of maize.

XX Claim 1; Page 104-107; 138pp; English.

XX This is the nucleotide sequence of the maize gene d111 (d11). To

XX illustrate the role of the d11 locus in starch biosynthesis, a transposon

XX tagging strategy was used to isolate the gene and describe its

XX polypeptide product. The invention reports tagging of the d11 locus with

XX Mu transposon, cloning and characterization of a portion of the gene, and

XX isolation of a near full-length cDNA (the present sequence). The amino acid

XX sequence (see AAY06199) deduced from this cDNA indicates that d11 codes

XX for a 186 kDa protein extremely similar to potato tuber starch synthase

XX SSI1. Its expression pattern indicates that d11 codes for SSI1 of maize

XX endosperm. The d11 product contains unique sequence features in its N-

XX terminus that may mediate direct interactions with other starch

XX biosynthetic enzymes. Mutations within the maize SSI1 gene affect

XX multiple aspects of starch biosynthesis by disrupting an enzyme complex

XX containing starch synthase(s), starch branching enzyme(s) and possibly

XX starch debranching enzyme(s). The isolated cDNA can be used to provide an

XX enzyme with which to regulate the production of starch, and with which to

XX produce altered or novel forms of starch, e.g. in transgenic plants.

XX Expression of d11 in bacteria and yeasts also modifies glycogen

XX production. Claimed expression vectors comprise the cDNA or fragments of

XX it that code for functional portions of d11

XX Sequence 6027 BP; 1800 A; 1177 C; 1441 G; 1609 T; 0 U; 0 Other;

XX Query Match 4.8%; Score 66.4; DB 2; Length 6027;

XX Best Local Similarity 77.5%; Pred. No. 0.00028;

XX Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;

XX DB 792 TTTAAGCTAGTTGGCAACCTGTTCTTTCAAGATTGTTTTCAA 851

XX DB 5606 TCTACGGCTAGTTGGCAACCTGTTCTTTCAAGATTGTTTTCAA 5548

XX QY 852 TTAGTTTATTTCTGTTTAAATAGAAACCTAGTGAATAATAGTTCAGACTAG 911

XX DB 5547 TTAGTTTATTTCTGTTTAAATAGAAACCTAGTGAATAATAGTTCAGACTAG 5488

XX RESULT 12

XX ABX09935/c

XX ID ABX09935 standard; DNA; 6027 BP.

XX AC ABX09935;

XX DT 17-FEB-2003 (first entry)

XX DE DNA encoding maize starch synthase III (D11).

XX Starch; starch synthase; glucan association domain; GLASS; linker domain;

XX LINKR; glucosyl transferase domain; GLYTR; C-terminal end; CTEND;

XX granule bound starch synthase; GBSS; morphology; retrogradation;

XX waterbinding; swelling potential; gene; ds.

XX Zea mays.

XX WO200279410-A2.

XX PN 10-OCT-2002.

XX PD 29-MAR-2002; 2002WO-US009574.

XX PF

XX 30-MAR-2001; 2001US-0279720P.
XX (BADI) BASF PLANT SCI GMBH.
XX
XX Commuri P, Keeling PL, Ramirez N, McKean A, Gao Z, Guan H;
XX WPI; 2003-040678/03.
XX
XX
XX New DNA encoding fusion protein consisting of 4 different functional
XX domains selected from glucan association domain, linker domain, glucosyl
XX transferase domain, and C-terminal end, useful for producing modified
XX starches.
XX
XX Claim 33; Page 225-227; 265pp; English.
XX
XX The invention describes an isolated DNA molecule encoding a fusion
XX protein consisting of 4 different functional domains selected from glucan
XX association domain (GLASS), linker domain (LINKR), glucosyl transferase
XX domain (GLYTR), and C-terminal end (CTEND) which are operably linked to
XX one another. The DNA molecule is useful for expressing in plants
XX polypeptides including starch synthase enzymes as fusion proteins with
XX improved affinity to starch and modified catalytic capabilities and to
XX the in vivo and in vitro synthesis of glucan chains of modified lengths
XX as compared to plants producing native starch or starch produced with
XX native starch synthases. Expression of the starch synthase fusion
XX proteins along with granule bound starch synthase (GBSS) will lead to a
XX modified starch having an altered or improved morphology, retrogradation,
XX waterbinding, or swelling potential of the granules, gel strength,
XX adhesiveness, cohesiveness, hardness, elasticity, increased or decreased
XX granule size, degree of branching, crystallinity, degree of cross-
XX linking, and increased or decreased glucan chain lengths. This sequence
XX encodes a starch synthase used in the invention
XX
SQ Sequence 6027 BP, 1800 A, 1177 C, 1441 G, 1609 T; 0 U, 0 Other;
XX
Query Match 4.8%; Score 66.4; DB 8; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00028;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
QY 792 TTATAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAGAAAAA 851
DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTCCAGAAAAA 5548
XX
QY 852 TTAGTTATTTTCTTTATAAATAGAAAAACCTAGAAAAATAGTTGCGAGCTAG 911
DB 5547 TTAGTTATTTTCTTTATTTGAAAAAATTTGAAATCTTTGAAAAATAGAGTTCACTACTAG 5488
XX
RESULT 13
ADK12106/c
ID ADK12106 standard; cDNA; 6027 BP.
XX
AC ADK12106;
XX
DT 03-JUN-2004 (first entry)
XX
DE cDNA encoding maize starch synthase, DUI.
XX
XX Maize; dull 1; DUI; starch synthase;
XX alpha-1,4-glycosyltransferase catalytic activity; starch production;
XX plant; gene; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 120..514
FT /tag= a
FT /product= "DUI"
XX
PN US2004049810-A1.
XX
PD 11-MAR-2004.

XX 05-AUG-2003; 2003US-00634262.
XX
XX 12-NOV-1997; 97US-00968467.
XX
XX 12-MAY-2000; 2000US-00554467.
XX
XX (MYER/) MYERS A M.
XX (JAME/) JAMES M G.
XX
XX Myers AM, James MG;
XX
XX WPI; 2004-238526/22.
XX
XX P-PsDB; ADK12117.
XX
XX
XX New nucleic acid designated dull1, encoding a starch synthase, useful in
XX starch production.
XX
XX Claim 1; SEQ ID NO 1; 58pp; English.
XX
XX The present invention relates to the isolation of a maize gene, dull 1
XX (DUI), and the polypeptide it encodes. The DUI polypeptide has starch
XX synthase activity, and comprises an N-terminal arm region, a C-terminal
XX catalytic region, and a region of about 900 amino acids terminating with
XX the catalytic region. The C-terminal catalytic region has a catalytic
XX domain comprising alpha-1,4-glycosyltransferase catalytic activity. The
XX dull polynucleotide sequence is useful in producing starch e.g. from a
XX transgenic plant or transfected cell. The present sequence encodes maize
XX DUI.
XX
SQ Sequence 6027 BP, 1800 A, 1177 C, 1441 G, 1609 T; 0 U, 0 Other;
XX
Query Match 4.8%; Score 66.4; DB 12; Length 6027;
Best Local Similarity 77.5%; Pred. No. 0.00028;
Matches 93; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
QY 792 TTATAGAGCTAGTTGGCAACCTGTTCTTTCAAGAAATTTGATTTTTCAGAAAAA 851
DB 5606 TCTACGGGCTAGTTGGGAACCCCATTT-TTCCAAGGATTTCCATTTTCCAGAAAAA 5548
XX
QY 852 TTAGTTATTTTCTTTATAAATAGAAAAACCTAGAAAAATAGTTGCGAGCTAG 911
DB 5547 TTAGTTATTTTCTTTATTTGAAAAAATTTGAAATCTTTGAAAAATAGAGTTCACTACTAG 5488
XX
RESULT 14
ABX35844
ID ABX35844 standard; cDNA; 439 BP.
XX
AC ABX35844;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #1009.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; IMF;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US200237139-A1.
XX
PD 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-015707P.
XX
XX 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARRE/) WARREN W C.


```

XX  Byate JC, Mathalagan N, Tao N, Warren WC;
XX  MPI; 2003-110599/10.
XX  New nucleic acid associated with lactation, and muscle and fat
XX  deposition, useful for genome mapping, gene identification and analysis,
XX  cattle breeding, or for genetically improving cattle.
XX  Claim 2; SEQ ID NO 1009; 245bp; English.
XX
XX  The invention relates to a purified nucleic acid molecule associated with
XX  lactation or muscle and fat deposition (designated LMP), derived from
XX  cattle, and the LMP nucleic acid can specifically hybridise to a second
XX  nucleic acid molecule comprising any of 1512 nucleotide sequences,
XX  appearing as ABX34836-ABX49947, or complements of them. Also included are
XX  (1) a transformed cell having a nucleic acid comprising an LMP nucleic
XX  acid linked to a promoter and a 3' non-translated sequence that
XX  functions in the cell to cause termination of transcription and addition
XX  of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX  (2) determining a level or pattern of a molecule in a bovine cell or
XX  tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX  of the 1512 nucleic acid sequences or its complement or fragment) with a
XX  complementary nucleic acid molecule obtained from the bovine cell or
XX  tissue, where hybridisation between the marker nucleic acid and the
XX  complementary nucleic acid permits the detection of the molecule; and (b)
XX  detecting the level or pattern of the complementary nucleic acid, where
XX  the detection of the complementary nucleic acid is predictive of the
XX  level or pattern of the molecule. The LMP nucleic acid is used for
XX  determining a level or pattern of a molecule in a bovine cell or tissue.
XX  It is useful for genome mapping, gene identification and analysis, cattle
XX  breeding, preparation of constructs for use in cattle gene expression, or
XX  for genetically improving cattle. The present sequence is one of the
XX  1512 bovine LMP EST (expressed sequence tag) nucleic acids. Note: The
XX  present sequence was not shown in the specification but was obtained in
XX  electronic format from the USPTO web site:
XX  seqdata.uspto.gov/sequence.html?DocId=20020137139
XX
XX  Sequence 439 BP; 45 A; 51 C; 56 G; 286 T; 0 U; 1 Other;
XX
XX  Query Match
XX  Best Local Similarity 54.7%; Score 66.2; DB 8; Length 439;
XX  Matches 134; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
XX
QY 626 TTCAGATTTTCTTTTCATCTGTTATTTGTTATGTTTATATACATTTCTCT 685
DB 182 TTCTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTATTTCTT 241
QY 686 CTTACATAGAGGATTTCTCCGATTTTATAAATGACTATTAAGTCATTTTATA 745
DB 242 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 301
QY 746 AGACGACGATGCTAGATTCGTTCAAAAATCTTGCATTTTTTAAAGACTAGT 805
DB 302 TGTGGGGGGGTGTTTTTTTTTTTTTTTTTTTTTTTCTTTTTTTTTTTTTTT 361
QY 806 TGGCAACCTGTTCTTTCAAGATTTGATTTTCAAAAAAATAGTATTTTCT 865
DB 362 TTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTATTTCTTATTTTTTTTTTT 421
QY 866 CTTTATA 872
DB 422 TTCTTTA 428
XX
RESULT 15
AA187279/c
ID AA187279 standard; cDNA; 346 BP.
XX
AC AA187279;
XX
DT 06-NOV-2001 (first entry)
XX

```

```

DE Human polynucleotide SEQ ID NO 7339.
XX
XX  Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX  vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX  tissue growth factor; immunomodulatory; cancer; leukaemia;
XX  nervous system disorders; arthritis; inflammation; ss.
XX
XX  Homo sapiens.
XX
XX  WO200164835-A2.
XX
XX  07-SEP-2001.
XX
XX  26-FEB-2001; 2001WO-US004927.
XX
XX  28-FEB-2000; 2000US-00515126.
XX
XX  18-MAY-2000; 2000US-00577409.
XX
XX  (HYSF-) HYSEQ INC.
XX
XX  Tang YT, Liu C, Drmanac RT;
XX
XX  MPI; 2001-514838/56.
XX
XX  P-PSDB; AA007348.
XX
XX  Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX  and treating e.g. leukemia, inflammation and immune disorders.
XX
XX  Claim 1; SEQ ID NO 7339; 1399pp + Sequence Listing; English.
XX
XX  The invention relates to human polynucleotides (AA179941-AA193841) and
XX  the encoded proteins (AA000010-AA013910) that exhibit activity elating to
XX  cytokine, cell proliferation or cell differentiation or which may induce
XX  production of other cytokines in other cell populations. The
XX  CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX  peptide therapy. The polypeptides have various cytokine-like activities,
XX  e.g. stem cell growth factor activity, haematopoiesis regulating
XX  activity, tissue growth factor activity, immunomodulatory activity and
XX  activin/inhibin activity and may be useful in the diagnosis and/or
XX  treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX  inflammation. Note: The sequence data for this patent did not form part
XX  of the printed specification, but was obtained in electronic format
XX  directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX  Sequence 346 BP; 188 A; 22 C; 16 G; 120 T; 0 U; 0 Other;
XX
XX  Query Match
XX  Best Local Similarity 4.7%; Score 65.6; DB 4; Length 346;
XX  Matches 143; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
XX
QY 446 AGATTTAAAAAATAACCAAGTAACTATTCACATCGAAAGCTATCATGTATTTAA 505
DB 334 ACAAAATGAATATATATTAATAAATGAAGATTAAATTTATATCATTAATAATTTTA 275
QY 506 GAAACATCATTAATAAACCAAGATCTCTTAAAAAAACAAGCATTTTCGAAAAGACAAT 565
DB 274 GACTTATATATATTAATAATTTTTTTTTTTTGAATAATATTTTAAACAAATAAAAA 215
QY 566 TATGTTACGTTTAAACATCTAGAGGACCAATTTATATGAAAGGTAAGCTATGACG 625
DB 214 TTTTAAATTAATAAATTAATAATTAATAATTAATAATTAATAATTTTATT 155
QY 626 TTCAGATTTTCTTTTCATCTGTTATTTGTTATGTTTATATACATTTTCTCT 685
DB 154 TATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 95
QY 686 CTTACATAGAGGATTTCTCCGATTTTAT 717
DB 94 TTTTTTTTTTTTTTTTTTTTTTTCTTTTTT 63
XX

```

Search completed: September 15, 2005, 21:33:03
 Job time : 679.892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 210.859 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1

Sequence: 1 ccctggcgcgcctcatgaataa.....cttttcgcaccaccaccatg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB. seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B COMB. seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB. seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB. seq.*
 - 5: /cgn2_6/ptodata/1/ina/PTUS COMB. seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1394	100.0	1394	3	US-08-880-499-1
2	1389.2	99.7	1394	3	US-08-880-499-2
3	66.4	4.8	6027	2	US-08-968-542C-1
4	66.4	4.8	6027	2	US-09-554-467A-1
5	58.6	4.2	2523	4	US-08-410-784A-3
6	57	4.1	1914	2	US-08-487-826B-13
7	55.8	4.0	279	4	US-09-313-294A-5397
8	55.6	4.0	612	4	US-09-902-540-1357
9	55.6	4.0	1394	3	US-08-880-499-1
10	55.6	4.0	1394	3	US-08-880-499-2
11	55.4	4.0	1055	4	US-09-806-708B-23
12	55.2	4.0	2614	4	US-09-004-056-1
13	55	3.9	336	4	US-09-640-173-53
14	55	3.9	336	4	US-09-713-550-53
15	55	3.9	336	4	US-09-825-294-53
16	55	3.9	336	4	US-09-970-866-53
17	54.6	3.9	5586	4	US-09-949-016-15129
18	53.8	3.9	307	4	US-09-313-294A-4743
19	53.8	3.9	6027	2	US-08-968-542C-1
20	53.8	3.9	6027	2	US-09-554-467A-1
21	52.2	3.7	1141	4	US-09-806-708B-22
22	51.8	3.7	1141	4	US-09-806-708B-22
23	51.6	3.7	134987	4	US-09-949-016-15348
24	51.6	3.7	134987	4	US-09-949-016-15349
25	51.6	3.7	134987	4	US-09-949-016-15350
26	51.6	3.7	134987	4	US-09-949-016-15507
27	51.6	3.7	134987	4	US-09-949-016-15508

C 28	51.6	3.7	134987	4	US-09-949-016-15509	Sequence 15509, A
C 29	51.4	3.7	50383	4	US-09-949-016-17600	Sequence 17600, A
C 30	51.4	3.7	129415	4	US-09-949-016-16997	Sequence 16997, A
C 31	51	3.7	601	4	US-09-949-016-156535	Sequence 156535, A
C 32	50.6	3.6	1039	4	US-09-902-540-1280	Sequence 1280, Ap
C 33	50.6	3.6	16573	4	US-09-949-016-14876	Sequence 14876, A
C 34	50.6	3.6	18773	4	US-09-949-016-14164	Sequence 14164, A
C 35	50.4	3.6	721	1	US-08-451-405A-2	Sequence 14164, A
C 36	50.2	3.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
C 37	50.2	3.6	317366	4	US-09-949-016-16001	Sequence 16001, A
C 38	50	3.6	1392	4	US-09-257-584-1	Sequence 1, Appl1
C 39	49.8	3.6	2435	3	US-09-306-593-1	Sequence 1, Appl1
C 40	49.8	3.6	231129	4	US-09-949-016-16110	Sequence 16110, A
C 41	49.8	3.6	266283	4	US-09-949-016-11934	Sequence 11934, A
C 42	49.6	3.6	640681	4	US-09-790-988-1	Sequence 1, Appl1
C 43	49.2	3.5	126176	4	US-09-949-016-16137	Sequence 16137, A
C 44	49.2	3.5	126176	4	US-09-949-016-16138	Sequence 16138, A
C 45	47.8	3.4	601	4	US-09-949-016-25787	Sequence 25787, A

ALIGNMENTS

RESULT 1

US-08-880-499-1

Sequence 1, Application US/08880499

Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Alberson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garraat W.

APPLICANT: Huffman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: CONCURRENTLY HERewith

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELEPHONE/COMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 1394; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATCCGTTTCTTAAAGGATCC 60
 DB 1 CCATGGTGTCTATGAAAAAGATGAGTACATGTGTCTATCCGTTTCTTAAAGGATCC 60
 QY 61 CTTCCTTCGCTTAATTAAGTGAATCGGGGTTTACAAAAAATTCCACGGGTGATGAT 120
 DB 61 CTTCCTTCGCTTAATTAAGTGAATCGGGGTTTACAAAAAATTCCACGGGTGATGAT 120
 QY 121 CTCGATGTTCCACTTCCCACTCCGCTTGCACTTCTTGATGTGCGTGGTCCCAT 180
 DB 121 CTCGATGTTCCACTTCCCACTCCGCTTGCACTTCTTGATGTGCGTGGTCCCAT 180
 QY 181 CTGACCGAGGCCCATGAGACACCTTTCGAGACACCATCAAGGGCTTTCGATGGCCCA 240
 DB 181 CTGACCGAGGCCCATGAGACACCTTTCGAGACACCATCAAGGGCTTTCGATGGCCCA 240
 QY 241 CGAGACGTATCGGGTCGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
 DB 241 CGAGACGTATCGGGTCGTGTGATCCAGGGGATATATGTCGCCCAATGTCACCTATA 300
 QY 301 TTATATCTCTTGAATATTAATTTTGGAAAAATTAACAATCTATCTTTGTGTA 360
 DB 301 TTATATCTCTTGAATATTAATTTTGGAAAAATTAACAATCTATCTTTGTGTA 360
 QY 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAATG 420
 DB 361 GGGCTCAGCATGATTTTCGTTAGGGCCAGAAAATGCGAGGACGAGCCATGCTAATG 420
 QY 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACAAAGTATCTAATCAGT 480
 DB 421 TCCACTATTTGGCACTACCCAGAACAGATTTAAAAAATTAACAAAGTATCTAATCAGT 480
 QY 481 CGAAGCTATCATGTATGTTTAAAGAAACATATTTAAACGACATCTCTTAAAAAA 540
 DB 481 CGAAGCTATCATGTATGTTTAAAGAAACATATTTAAACGACATCTCTTAAAAAA 540
 QY 541 CAAGCATATTTGAAAAAGACAAATATGTATACAGTTTACAAACATCTAAGGCGACAA 600
 DB 541 CAAGCATATTTGAAAAAGACAAATATGTATACAGTTTACAAACATCTAAGGCGACAA 600
 QY 601 TTATATCGAAAGGTAAAGTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTGTT 660
 DB 601 TTATATCGAAAGGTAAAGTATGACGTTCAAGATTTTCTTTTCAATCTGTTATTTGTT 660
 QY 661 ATGTTTTTATATACATTTCTCTCTTACATAGATGATTTCTTCGATTTTAA 720
 DB 661 ATGTTTTTATATACATTTCTCTCTTACATAGATGATTTCTTCGATTTTAA 720
 QY 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGCTAGATTCGTTCAAAAAATC 780
 DB 721 ATGACTATAAAGTCATTTTATATAAGAGCAGCATGCTAGATTCGTTCAAAAAATC 780
 QY 781 TTTCTGATTTTTTAAGAGTACTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 DB 781 TTTCTGATTTTTTAAGAGTACTTTGGCAACCTGTTTCTTCAAGAAATTTGATTTT 840
 QY 841 TTCAAAAAAAATAGTTTATTTTCTCTTATAAATAGAAAAACATAGAAAAATGAGT 900
 DB 841 TTCAAAAAAAATAGTTTATTTTCTCTTATAAATAGAAAAACATAGAAAAATGAGT 900
 QY 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATTAATTGG 960
 DB 901 TGGCAGACTAGCCCTAGAAATGTTTCCCAATAAATTAACATCACTGTATATTAATTGG 960
 QY 961 GCCAGGCCCATTAATTTTAAACCGAACTGAATTCGAGCGAAACCAATCTGAGCTAT 1020
 DB 961 GCCAGGCCCATTAATTTTAAACCGAACTGAATTCGAGCGAAACCAATCTGAGCTAT 1020
 QY 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
 DB 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTTAAAGTCAATGCTCC 1080
 QY 1081 TGAGATGTGGGTTTGGCAACGATAGCACCGTAATCATAGTCAATAGGTGCTTACGTCA 1140

DB 1081 TGAGATGTGGGTTTGGCAACGATAGCACCGTAATCATAGTCAATAGGTGCTTACGTCA 1140
 QY 1141 GGTTCGGCAGCTCTCGTGTATCTCATGAGCATATACATCTGTTCAACCGTTGCTC 1200
 DB 1141 GGTTCGGCAGCTCTCGTGTATCTCATGAGCATATACATCTGTTCAACCGTTGCTC 1200
 QY 1201 TTGTTCCATGCTGCAAGCCTTTCGATTTCTGAACCAAGGATPACTACTCCCAACAT 1260
 DB 1201 TTGTTCCATGCTGCAAGCCTTTCGATTTCTGAACCAAGGATPACTACTCCCAACAT 1260
 QY 1261 CCATCTTACTCATGCACTTCCATGCAACAGCACATATGTTTCTGAACCAATCCATT 1320
 DB 1261 CCATCTTACTCATGCACTTCCATGCAACAGCACATATGTTTCTGAACCAATCCATT 1320
 QY 1321 AAAGATCACAACGCTAGCGTTTCCCGTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 DB 1321 AAAGATCACAACGCTAGCGTTTCCCGTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
 QY 1381 CGTCCACGACGATG 1394
 DB 1381 CGTCCACGACGATG 1394

RESULT 2
 US-08-880-499-2
 ; Sequence 2, Application US/08880499
 ; Patent No. 6037523
 ; GENERAL INFORMATION:
 ; APPLICANT: Albertson, Marc C.
 ; APPLICANT: Fox, Tim W.
 ; APPLICANT: Carl, Garnat W.
 ; APPLICANT: Huffman, Gary A.
 ; APPLICANT: Kendall, Timmy L.
 ; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
 ; NUMBER OF INVENTIONS: 2
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
 ; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
 ; STREET: Box 1000
 ; CITY: Johnston
 ; STATE: Iowa
 ; COUNTRY: USA
 ; ZIP: 50131
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/880,499
 ; FILING DATE: CONCURRENTLY HERewith
 ; CLASSIFICATION: 800
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sweeney, Patricia A.
 ; REGISTRATION NUMBER: 32,733
 ; REFERENCE/DOCKET NUMBER: 0578
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (515) 248-4800
 ; TELEFAX: (515) 248-4844
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1394 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-880-499-2
 Query Match 99.7%; Score 1389.2; DB 3; Length 1394;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```

Qy 1 CCATGGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Db 1 CCATGGTCTCTATGAAAAAGATGATACATGTGTCTATATCCGTTTTCTTAGGGTCC 60
Qy 61 CTTCTCTGCTTATTAAGTGAATCGGGGTTTACAAAAAATTCCAGGGGGCATGTAT 120
Db 61 CTTCTCTGCTTATTAAGTGAATCGGGGTTTACAAAAAATTCCAGGGGGCATGTAT 120
Qy 121 CTCATGTTTCACTTCTCCCACTCGCGTTCACATTTCTGGATTCGGTGGTCCAT 180
Db 121 CTCATGTTTCACTTCTCCCACTCGCGTTCACATTTCTGGATTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCCATCAGACACTTTTCGGGACACCCATCAAGGGCCTTTTCGATGGCCA 240
Db 181 CTGACCGAGGCCCATCAGACACTTTTCGGGACACCCATCAAGGGCCTTTTCGATGGCCA 240
Qy 241 CGAGACGTATCGGGGTGATGATCGAGGGGATATATGTCCCCCAATCGTCACTTAT 300
Db 241 CGAGACGTATCGGGGTGATGATCGAGGGGATATATGTCCCCCAATCGTCACTTAT 300
Qy 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATAACAATTAATTAATTTGTAT 360
Db 301 TTATTAATCTTATGATATTAATTAATTTTGGAAAAATAACAATTAATTAATTTGTAT 360
Qy 361 GGGCCTCAGCATGATTTTTCGCTTAGGGCCAGAAAATCGAGAGACAGCCATGTCTAGTG 420
Db 361 GGGCCTCAGCATGATTTTTCGCTTAGGGCCAGAAAATCGAGAGACAGCCATGTCTAGTG 420
Qy 421 TCCACTATTGGGCTACCGAGAACAGATTTTAAAAAATPACCAAGTAATCAATCACT 480
Db 421 TCCACTATTGGGCTACCGAGAACAGATTTTAAAAAATPACCAAGTAATCAATCACT 480
Qy 481 CGAAAGCTATCATGTAATGTTTAAAGAAATCTATTAATAACACGATCCTTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAATGTTTAAAGAAATCTATTAATAACACGATCCTTTAAAAAA 540
Qy 541 CAAGCATATTTGAAAGAGACAAATATGTATACATTTCACAAATCTTAAGAGCGAATA 600
Db 541 CAAGCATATTTGAAAGAGACAAATATGTATACATTTCACAAATCTTAAGAGCGAATA 600
Qy 601 TTATATCGAAGGTAGCTATGCGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
Db 601 TTATATCGAAGGTAGCTATGCGTTCAGATTTTCTTTTCATCTTGTATTTTGT 660
Qy 661 ATTTGTTTTATATACATTTTCTTCTCAATAGAGATTTTCTTCGATTTTATAA 720
Db 661 ATTTGTTTTATATACATTTTCTTCTCAATAGAGATTTTCTTCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATTAAGACACGATGCGTAGATTCGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATTAAGACACGATGCGTAGATTCGTTCAAAAATC 780
Qy 781 TTTTGAATTTTAAAGCTAGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Db 781 TTTTGAATTTTAAAGCTAGTGGCAACCTGTTTCTTCAAGAAATTTGATTT 840
Qy 841 TTTCAAAAAAATAGTTTATTTTCTTCTTAAATAGAAAACATTTAGAAAATAGAT 900
Db 841 TTTCAAAAAAATAGTTTATTTTCTTCTTAAATAGAAAACATTTAGAAAATAGAT 900
Qy 901 TGGCAGACTAGCCCTAGAAATGTTTTCGAATTAATCAATCACTGTGTATTAATTTTG 960
Db 901 TGGCAGACTAGCCCTAGAAATGTTTTCGAATTAATCAATCACTGTGTATTAATTTTG 960
Qy 961 GCGAGCCCATAAATTAATTTAAACGAACTGAATCGAGGAAACCAATCGACTAT 1020
Db 961 GCGAGCCCATAAATTAATTTAAACGAACTGAATCGAGGAAACCAATCGACTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAATCAATGTCCC 1080
Db 1021 TTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCACTTTTAAATCAATGTCCC 1080

```

```

Qy 1081 TGAATGTGCGGTTTGGCAACATAGCCCACTTAATCATAGTCAATAGTGGCTACGTCA 1140
Db 1081 TGAATGTGCGGTTTGGCAACATAGCCCACTTAATCATAGTCAATAGTGGCTACGTCA 1140
Qy 1141 GGTTCGAGCTCTCGTGCATCTCAGATGAGATACATGCTTGTTCACCGTTTCGTC 1200
Db 1141 GGTTCGAGCTCTCGTGCATCTCAGATGAGATACATGCTTGTTCACCGTTTCGTC 1200
Qy 1201 TTGTTTCATGTCGAAGCCTTTCCTATTCGAACCAAGAGATACCTACCCCAACAT 1260
Db 1201 TTGTTTCATGTCGAAGCCTTTCCTATTCGAACCAAGAGATACCTACCCCAACAT 1260
Qy 1261 CCATCTTACTATGCAATTCATGCAATGCAAAACAGCACATATGTTTCTGAACCAATTCAT 1320
Db 1261 CCATCTTACTATGCAATTCATGCAATGCAAAACAGCACATATGTTTCTGAACCAATTCAT 1320
Qy 1321 AAAAGATCAACAGGAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAAGATCAACAGGAGGTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCCACCAACATG 1394
Db 1381 CGTCCACCAACATG 1394
Qy 1394 CGTCCACCAACATG 1394
Db 1394 CGTCCACCAACATG 1394

RESULT 3
US-08-542C-1/c
; Sequence 1, Application US/08968542C
; Patent No. 5981728
; GENERAL INFORMATION:
; APPLICANT: Myers, et al.
; TITLE OF INVENTION: dulli Codes For A No. 5981728el Starch
; TITLE OF INVENTION: Synchase
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: McGregor & Adler, LLP
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word 6.0.1 for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/968,542C
; FILING DATE: No. 5981728ember 12, 1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D6036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6027 bp
; TYPE: nucleic acid
; STRANDEDNESS: double-stranded
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: cDNA to mRNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: maize
; TISSUE TYPE: endosperm

```



```

; IMMEDIATE SOURCE:
; LIBRARY: maize endosperm cDNA library in
; LIBRARY: (gt11
; CLONE: pmG6Aa; pmG6-2M
; CLONE: pmGf10;
US-08-968-542C-1

```

Query Match	4.8%;	Score 66.4;	DB 2;	Length 6027;
Beat Local Similarity	77.5%;	Pred. No. 2.2e-06;		
Matches 93;	Conservative 0;	Mismatches 26;	Indels 1;	Gaps 1;

Qy 792 TTATAGAGCTAGTTGGCAACCCGTTCTTTCAAGAATTTGATTTTTCAAAAA 851

Db 5606 TCTAGCGGCTAGTTGGCAACCCCATTT - TTCAGAGGATTTCCATTTTCCAGAAAA 5548

Qy	852	TTAGTTTATTTCTCTTATATAAATAGAAAACACTAGAAAATGAGTCCGACTAG	912
Db	5547	TTAGTTTATTTTCAITGGAAAAATGAAATCTCTTGGAAAAATAGAGTTCACATACTAG	5488

RESULT 4
US-09-554-467A-1/C

```

: GENERAL INFORMATION:
: APPLICANT: Myers, Alan W.
: APPLICANT: James, Martha G.
: TITLE OF INVENTION: d,lil Coding for a No. 6639125e1 Starch Synthase and Uses
: TITLE OF INVENTION: Thereof
: FILE REFERENCE: D6036PCT
: CURRENT APPLICATION NUMBER: US/09/554,467A
: CURRENT FILING DATE: 2000-05-12
: PRIOR APPLICATION NUMBER: PCT/US98/24225
: PRIOR FILING DATE: 1998-11-12
: PRIOR APPLICATION NUMBER: US 08/063,102
: PRIOR FILING DATE: 1997-11-12
: NUMBER OF SEQ ID NOS: 37

```

```

; OTHER INFORMATION: cDNA sequence corresponding to the gene encoding the
; OTHER INFORMATION: starch synthase enzyme D11.
US-09-554-467A-1

```

Query Match	4.8%	Score 66.4	DB 4	Length 6027
Best Local Similarity	77.5%	Pred. NO. 2.2e-06		
Matches 93; Conservative	0	Mismatches 26	Indels 1	Gaps 1

QY 852 TTAGTTTATTTCTCTTATTAATAAGAAAACACTTGAATAAGATTGCCAGACTTG 911

Db 5547 TTGGTTTATTTTTCATTGGAAAAATGGAATCTCTGGAAAAATAGAGTTCCACTACTTG 5488

RESULT 5
US-08-410-784A-3/C
; Sequence 3, Application US/08410784A

GENERAL INFORMATION: ALAN M.
APPLICANT: MYERS, ALAN M.
APPLICANT: JAMES, MARTHA G
TITLE OF INVENTION: ISOLATION OF SU1, A STARCH DEBRANCHING
TITLE OF INVENTION: ENZYME, THE PRODUCT OF THE MAIZE GENE
TITLE OF INVENTION: SUGARY 1
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: weingarten, schurgin, Gagnebin and Hayes LLP
STREET: Ten Post Office Square
CITY: Boston

```

; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READAB

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:

Query Match	4.2%;	Score 58.6;	DB 2;	Length 2523;
Best Local Similarity	71.4%;	Pred. No. 0.00013;		
Matches 105; Conservative	0;	Mismatches 39;	Indels 3;	Gaps 2

```

Oy      774  AAAAAATCTTCTGATTTTAAAAAGAGCTAGTTGGCAACCCGTTCCTTTCAAAAGATTT 833
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      279  ACATATACATATACATATAAGCTCTAGTTGACAATCCATTTTATCAAAG--TTT 222

```

Qy 834 TGAATTTTCAAAAAAAAAATTAGTTATTTTCTCTTTATAAATAGAAAAACACTTAGAAAA 8933

Db 221 TACATTTTTCAAATTAATTAGTTATTTTCTCTTGA-AAAATAGGAATTTTCAGAAAA 1631

QY	894	ATAGAGTTGCCAGACTAGCCCTAGAT	920
Db	162	ATAGAGTTTACAACACTAGTCTTAAAT	136

RESULT 6
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B

GENERAL INFORMATION:
 APPLICANT: Sim, Kim L.
 APPLICANT: Chitnis, Chetan
 APPLICANT: Miller, Louis H.
 APPLICANT: Peterson, David S.
 APPLICANT: Su, Xin-zhaun
 APPLICANT: Wellem, Thomas E.
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Knobbe Martens Olson & Bear
 STREET: 620 Newport Center Drive 16th Floor
 CITY: Newport Beach
 STATE: California
 COUNTRY: US

ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,825B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
US-08-487-825B-13

Query Match 4.1%; Score 57; DB 2; Length 19124;
Best Local Similarity 50.2%; Pred. No. 0.00073;

Matches 141; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 614 TAAGCTATGACGCTTCAATTTCTTTCTTCTGTTATTTGTTATGTTTATAT 673
DB 15677 TAATGTTTTTTTTTTCTTCTTTGTTGTTTATTTTAAATCATTTTATAT 15618
QY 674 ACATTTCTCTCTTCAATAGAGTATTTCTCGATTTTAAAGCATATAAGT 733
DB 15617 AAAATTTTTTTTTTAAATTTTGTATATCTTTTCAATTTTATCTATCAAAAATTA 15558
QY 734 CATTTTATATTAAGACGACATGTCGTAGATTCGTCGTAATAATCTTCTGATTTTT 793
DB 15557 TATTTATATATATTTTATATTTTAAATTTTCTCTTTTATTTTATTTT 15498
QY 794 TAAAGCTAGCTGGGCAACCTGTTCTTCAAGATTTTGAATTTTCAAAAAAAT 853
DB 15497 TTTTATTTTAAATTTTATTTTATATTTTCAATTTTCTTTTCAATTTTAAATTT 15438
QY 854 AGTTATTTCTCTTATATAAATAGAAACACTAGAAAA 894
DB 15437 GTTTTATATTTCTTTTAAATTAATAATATATATAA 15397

RESULT 7

US-09-313-294A-5397
Sequence 5397, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Ragunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 5397
LENGTH: 279
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: mlec_feature

OTHER INFORMATION: Incyte ID No. 6476212 700350078H1
NAME/KEY: unsure
LOCATION: 10, 12, 185-186, 204, 253, 274, 278
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5397

Query Match 4.0%; Score 55.8; DB 4; Length 279;
Best Local Similarity 69.0%; Pred. No. 0.00029;
Matches 89; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 789 TTTTAAAGAGTAGTTTGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAA 848
DB 3 TTCTTAANGCTAGTTTGGAAACCCATTTTCCACGGTTTTCATTTTCCCAAGG 62
QY 849 AAATAGTTTATTTCTCTTAT-AAAATAGAAACCTTAGAAATAGAGTTGCCAG 907
DB 63 AAGTTAGAACATTTTCCCTTGGGAAATAGAAATCTTTGGGAAATGAGTTCCCAA 122
QY 908 CTAGCCCTA 916
DB 123 CTAGCCCTA 131

RESULT 8

US-09-902-540-1357/C
Sequence 1357, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 1357
LENGTH: 612
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(612)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 4.0%; Score 55.6; DB 4; Length 612;
Best Local Similarity 51.4%; Pred. No. 0.00043;
Matches 127; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 626 TTCAGATTTTCTTTTCAATCTGTATTTGTTATTTGTTTATATATTTCTCT 685
DB 560 TTTTATATTTTCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 501
QY 686 CTACATAGAGTATTTCTTCGATTTTATATAAGATTAAGTATTTATATA 745
DB 500 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 441
QY 746 AGAGCAGCATGTCGATGTTCTCGTCAAAAATCTTCTGATTTTAAAGCTAG 805
DB 440 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 381
QY 806 TGGCAACCTGTTCTTCAAGAAATTTGATTTTCAAAAAATAGTTATTTCT 865
DB 380 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 321
QY 866 CTTATA 872
DB 320 TATTTTA 314


```
RESULT 9
US-08-880-499-1/C
; Sequence 1, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
; APPLICANT: Alberson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnat W.
; APPLICANT: Hulftman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-1

Query Match      4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.0006;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY      761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTC 820
DB      950 TACACAGTATGTATTTATTTAGGAAACATTTAGGGCTAGCTGGCAACTTATTTT 891
QY      821 TTTCAGAAGATTTGATTTTTCAAAAAAATTAAGTTATTTCTCTTATAAATAGAA 880
DB      890 TCTAAGTGTCTTATTTATTAAGGAAATTAAGTAAATTTTGGAAAAATCAAAA 831
QY      881 AACCTTGAAGAAATAGAGTTGCGAGACTAGCCCTAGAAATGTTTCCCAATTAATACAA 940
DB      830 TTCTTTGAAGAAACAGGGTTCACAACTAGCTTTAAAAAATCAGAAAGATTTTGAA 771
QY      941 TCACGTGTGA 950
DB      770 CGAGAAATCTA 761
```

```
RESULT 10
US-08-880-499-2/C
; Sequence 2, Application US/08880499
; Patent No. 6037523
; GENERAL INFORMATION:
```

```
; APPLICANT: Alberson, Marc C.
; APPLICANT: Fox, Tim W.
; APPLICANT: Carl, Garnat W.
; APPLICANT: Hulftman, Gary A.
; APPLICANT: Kendall, Timmy L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
; TITLE OF INVENTION: AND METHOD OF USING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
; STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
; STREET: Box 1000
; CITY: Johnston
; STATE: Iowa
; COUNTRY: USA
; ZIP: 50131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/880,499
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Sweeney, Patricia A.
; REGISTRATION NUMBER: 32,733
; REFERENCE/DOCKET NUMBER: 0578
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (515) 248-4800
; TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-880-499-2

Query Match      4.0%; Score 55.6; DB 3; Length 1394;
Best Local Similarity 55.8%; Pred. No. 0.0006;
Matches 106; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY      761 TAGATTCTCGTCAAAAATCTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTTTC 820
DB      950 TACACAGTATGTATTTATTTAGGAAACATTTAGGGCTAGCTGGCAACTTATTTT 891
QY      821 TTTCAGAAGATTTGATTTTTCAAAAAAATTAAGTTATTTCTCTTATAAATAGAA 880
DB      890 TCTAAGTGTCTTATTTATTAAGGAAATTAAGTAAATTTTGGAAAAATCAAAA 831
QY      881 AACCTTGAAGAAATAGAGTTGCGAGACTAGCCCTAGAAATGTTTCCCAATTAATACAA 940
DB      830 TTCTTTGAAGAAACAGGGTTCACAACTAGCTTTAAAAAATCAGAAAGATTTTGAA 771
QY      941 TCACGTGTGA 950
DB      770 CGAGAAATCTA 761

RESULT 11
US-09-806-708B-23
; Sequence 23, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
```


LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCCTTTTCATCTGTTATTTGTTATGTTTTTAATACATTTCTCTCTTACA 691
DB 11 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTATTAATAAGCTAATAAGCTATTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130
QY 752 CGCATGTCGATGTCCTGCTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTCAGACCTAGACCTAGATGTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAAAGAAANATTAANNANNNANNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTTNNNNNNANGG 327

RESULT 14
US-09-713-550-53
Sequence 53, Application US/09713550
Patent No. 6617109
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolck, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.484C4
CURRENT APPLICATION NUMBER: US/09/713,550
CURRENT FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 205
SOFTWARE: FaSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCCTTTTCATCTGTTATTTGTTATGTTTTTAATACATTTCTCTCTTACA 691
DB 11 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTATTAATAAGCTAATAAGCTATTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130

QY 752 CGCATGTCGATGTCCTGCTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTCAGACCTAGACCTAGATGTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAAAGAAANATTAANNANNNANNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTTNNNNNNANGG 327

RESULT 15
US-09-825-294-53
Sequence 53, Application US/09825294
Patent No. 6710170
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolck, John A.
APPLICANT: Algate, Paul A.
APPLICANT: Fling, Steven P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY AND DIAGNOSIS OF OVARIAN CANCER
FILE REFERENCE: 210121.484C5
CURRENT APPLICATION NUMBER: US/09/825,294
CURRENT FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 215
SOFTWARE: FaSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(396)
OTHER INFORMATION: n = A,T,C or G
US-09-825-294-53

Query Match 3.9%; Score 55; DB 4; Length 396;
Best Local Similarity 43.8%; Pred. No. 0.00052;
Matches 139; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

QY 632 TTTTCCTTTTCATCTGTTATTTGTTATGTTTTTAATACATTTCTCTCTTACA 691
DB 11 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 70
QY 692 ATAGAGTGAATTTCTCCGATTTTATTAATAAGCTAATAAGCTATTTATATAAGACA 751
DB 71 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 130
QY 752 CGCATGTCGATGTCCTGCTCAAAATCTTCTGATTTTAAAGACTAGTTGGCAA 811
DB 131 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 190
QY 812 CCTGTTCTTCAAGAAATTTGATTTTCAAAAAAATAGTTATTTCTCTTAT 871
DB 191 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 250
QY 872 AAAATGAAAAACCTTAGAAAAATAGAGTTCAGACCTAGACCTAGATGTTCCCAAT 931
DB 251 AATTCANAAAAAGAAANAAAGAAANATTAANNANNNANNNNNNNNATNTNCTTNATA 310
QY 932 AAATTACATCACTGTG 948
DB 311 NTNNTTNNNNNNANGG 327

Search completed: September 15, 2005, 08:24:54
Job time : 214.859 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 07:10:11 ; Search time 907.859 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccatggtgctctatgaaaaa.....cttttcgctcacaccacatg 1394

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
18: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
19: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	1394	100.0	1394	20	US-10-713-381-1
2	1389.2	99.7	1394	20	US-10-713-381-2
3	238.8	17.1	225	20	US-10-713-381-9
4	146	10.5	158	20	US-10-713-381-3
5	80.8	5.8	1261	20	US-10-425-115-134230
6	78.8	5.7	1663	20	US-10-739-930-1857
7	74	5.3	320	18	US-10-425-114-22340

	8	74	5.3	624	18	US-10-425-114-16264	Sequence 16264, A
	9	73	5.2	1326	20	US-10-425-115-141826	Sequence 141826, A
	10	72.8	5.2	2445	18	US-10-425-114-32493	Sequence 32493, A
	11	72.8	5.2	2729	20	US-10-425-115-83293	Sequence 83293, A
	12	72.4	5.2	1203	20	US-10-425-115-51470	Sequence 51470, A
	13	71.8	5.2	928	18	US-10-425-114-17816	Sequence 17816, A
	14	71.8	5.2	1123	20	US-10-425-115-142853	Sequence 142853, A
	15	71.8	5.2	2537	18	US-10-425-114-11957	Sequence 31957, A
	16	71.8	5.2	2537	20	US-10-425-115-52216	Sequence 52216, A
	17	71.8	5.2	3607	18	US-10-425-114-31061	Sequence 31061, A
	18	71.8	5.2	3691	20	US-10-425-115-52219	Sequence 52219, A
	19	70.2	5.0	960	14	US-10-198-846-6381	Sequence 6381, A
	20	69.8	5.0	13400	21	US-10-343-477A-63	Sequence 63, Appl
	21	69.6	5.0	724	20	US-10-425-115-87756	Sequence 87756, A
	22	69.2	5.0	527	18	US-10-425-114-4041	Sequence 4041, Ap
	23	68.8	4.9	610	20	US-10-425-115-47165	Sequence 47165, A
	24	68.4	4.9	1215	18	US-10-425-114-24656	Sequence 24656, A
	25	68.4	4.9	2863	20	US-10-425-115-75310	Sequence 75310, A
	26	67.8	4.9	512	20	US-10-425-115-49781	Sequence 49781, A
	27	67.8	4.9	928	18	US-10-425-114-17816	Sequence 17816, A
	28	67.8	4.9	1123	20	US-10-425-115-142853	Sequence 142853, A
	29	67.4	4.8	1260	18	US-10-425-114-30881	Sequence 30881, A
	30	67.4	4.8	1411	20	US-10-425-115-149304	Sequence 149304, A
	31	67.2	4.8	2232	18	US-10-425-114-16607	Sequence 16607, A
	32	67.2	4.8	2249	18	US-10-425-114-20264	Sequence 20264, A
	33	66.4	4.8	6027	18	US-10-634-262-1	Sequence 1, Appli
	34	66.4	4.8	6027	19	US-10-109-048-1145	Sequence 1145, Ap
	35	66.4	4.8	6051	20	US-10-425-115-41589	Sequence 41589, A
	36	66.2	4.7	439	9	US-09-960-352-1009	Sequence 1009, Ap
	37	66	4.7	1376	18	US-10-425-114-2073	Sequence 2073, Ap
	38	65.6	4.7	721	20	US-10-425-115-177935	Sequence 177935, Ap
	39	65.6	4.7	2274	18	US-10-425-114-4777	Sequence 777, App
	40	65.6	4.7	2274	18	US-10-425-114-830	Sequence 830, App
	41	65.6	4.7	2632	20	US-10-425-115-162957	Sequence 162957, A
	42	65.6	4.7	8056	20	US-10-473-126-1386	Sequence 386, App
	43	65	4.7	1364	20	US-10-425-115-150820	Sequence 150820, A
	44	64.8	4.6	833	20	US-10-425-115-30824	Sequence 30824, A
	45	64.8	4.6	1838	20	US-10-425-115-21012	Sequence 21012, A

ALIGNMENTS

RESULT 1
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNMAN, CARL W.
; APPLICANT: HUFFMAN, GARY L.
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 1394; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 116-294;
Matches 1394; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CcATGtGtCtCtATGAAAGATAGTACATGTGTCTATTCGTTTCTTAGGCTC 60

Db 1 CCAATGCTCTATGAAAAAGATGAGTACATGTCATATATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTTCGCTTATTAATGACTGAAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCGCTTATTAATGACTGAAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTCGGTGCACTTCTTGGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTCGGTGCACTTCTTGGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCATCGACACCTTTGGGACACCCATCAAGGACCTTTGGATGGCCA 240
Db 181 CTGACCGAGGCCATCGACACCTTTGGGACACCCATCAAGGACCTTTGGATGGCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACCTATA 300
Qy 301 TTATTTCTTTAGATATTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTTCTTTAGATATTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Qy 361 GGGCTTCAGATGATTTTGGCTTAAAGGCCAGAAATGCGAGACGACCATGTCTAGTG 420
Db 361 GGGCTTCAGATGATTTTGGCTTAAAGGCCAGAAATGCGAGACGACCATGTCTAGTG 420
Qy 421 TCCATTTTGGCTTAAAGGCCAGAAATGCGAGACGACCATGTCTAGTG 480
Db 421 TCCATTTTGGCTTAAAGGCCAGAAATGCGAGACGACCATGTCTAGTG 480
Qy 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCCTTTAAATA 540
Db 481 CGAAGCTATCATGTAATGTTTAAAGAAACATCTATTAACACGATCCTTTAAATA 540
Qy 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTCAACATCTAAAGCGACAA 600
Db 541 CAAGCATATTTGCAAGAGACAAATTAATGTTACAGTTTCAACATCTAAAGCGACAA 600
Qy 601 TTTATTCGAAAGGTAAGCTATGACGTTGCAATTTTCTTTTCAATCTTTGTT 660
Db 601 TTTATTCGAAAGGTAAGCTATGACGTTGCAATTTTCTTTTCAATCTTTGTT 660
Qy 661 ATTGTTTTTATATACATTTTCTTCTTCAATAGATGATTTCTTCCGATTTTATAA 720
Db 661 ATTGTTTTTATATACATTTTCTTCTTCAATAGATGATTTCTTCCGATTTTATAA 720
Qy 721 ATGACTATTAAGTCAATTTTATTAAGAGACGACGATGCTGATCTCGTTCAAAAATC 780
Db 721 ATGACTATTAAGTCAATTTTATTAAGAGACGACGATGCTGATCTCGTTCAAAAATC 780
Qy 781 TTTCTGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTAAAGCTATGTTGGCAACCTGTTCTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATTAGAAATAATGAGT 900
Db 841 TTCAAAAAAATAGTTTATTTCTCTTATTAATAAGAAACATTAGAAATAATGAGT 900
Qy 901 TGCAGACTAGCCCTGAAATGTTTCCCAATAAATCACTGTGTAATTAATTTG 960
Db 901 TGCAGACTAGCCCTGAAATGTTTCCCAATAAATCACTGTGTAATTAATTTG 960
Qy 961 GCGAGGCCATAATTAATTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACCTAT 1020
Db 961 GCGAGGCCATAATTAATTTAAACGAAACTGAAATCGAGCGAAACCAATCTGACCTAT 1020
Qy 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATCACTTTAAGTCACTTGTCC 1080
Db 1021 TTCTCTAGATTAGTAAAGGAGAGAGAGAGAAATCACTTTAAGTCACTTGTCC 1080
Qy 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTATCATAGTCAATGAGTGCCTAGCTCA 1140
Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTATCATAGTCAATGAGTGCCTAGCTCA 1140

Db 1081 TGAATGTGCGGTTTGGCAACGATAGCCAGTATCATAGTCAATGAGTGCCTAGCTCA 1140
Qy 1141 GGTGGGAGCTCTGTGTCATCTACATGAGATACATGCTTTTAAACCGTTCGTC 1200
Db 1141 GGTGGGAGCTCTGTGTCATCTACATGAGATACATGCTTTTAAACCGTTCGTC 1200
Qy 1201 TTGTTTCATGCTTCAGGCTTCTCTATTTCTGAACCAAGATATCTACTCCAAACAT 1260
Db 1201 TTGTTTCATGCTTCAGGCTTCTCTATTTCTGAACCAAGATATCTACTCCAAACAT 1260
Qy 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGAACCAATTCAT 1320
Db 1261 CCATCTTACTCATGCAACTTCCATGCAACGACATATGTTTCTGAACCAATTCAT 1320
Qy 1321 AAGATTCACACGTAAGTTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAGATTCACACGTAAGTTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCCACGACATG 1394
Db 1381 CGTCCACGACATG 1394

RESULT 2
US-10-713-381-2
; Sequence 2. Application us/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; CURRENT FILING DATE: 2003-11-14
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 99.7%; Score 1389.2; DB 20; Length 1394;
Best Local Similarity 99.8%; Pred. No. 1.2e-293;
Matches 1391; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCATGCTCTATGAAAAAGATGAGTACATGTCATATATCCGTTTCTTAGGGTCC 60
Db 1 CCATGCTCTCTATGAAAAAGATGAGTACATGTCATATATCCGTTTCTTAGGGTCC 60
Qy 61 CTTCTTCGCTTATTAATGACTGAAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT 120
Db 61 CTTCTTCGCTTATTAATGACTGAAATCGGGGTTACAAAAAATTCCACGGGTGCATGAT 120
Qy 121 CTCATGTTCCACTTCTCCACCTCGGTGCACTTCTTGGATGTCGGTGGTCCAT 180
Db 121 CTCATGTTCCACTTCTCCACCTCGGTGCACTTCTTGGATGTCGGTGGTCCAT 180
Qy 181 CTGACCGAGGCCATCGACACCTTTGGGACACCCATCAAGGACCTTTGGATGGCCA 240
Db 181 CTGACCGAGGCCATCGACACCTTTGGGACACCCATCAAGGACCTTTGGATGGCCA 240
Qy 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACCTATA 300
Db 241 CGAGAGCTATCGGGTGTGTGATCCAGGGGATATATGTCCCAATGTCACCTATA 300
Qy 301 TTATTTCTTTAGATATTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360
Db 301 TTATTTCTTTAGATATTTATTTTGAAGAAAAATACAACTTATCTTTGTGTA 360

Db 301 TTATATCTTTAGATATTTATTTATTTTGGAAAAATTAACAACTTAATCTTTGTGTA 360
Qy 361 GGGCCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATCGAGGACGACCATGCTAGTG 420
Db 361 GGGCCCTCAGCATAGATTTTCGCTTAGGGCCAGAAATCGAGGACGACCATGCTAGTG 420
Qy 421 TCCACTATTGGCACTACCCAGAACAAATTTAAAAATAACCAAGTAATTAATCCACT 480
Db 421 TCCACTATTGGCACTACCCAGAACAAATTTAAAAATAACCAAGTAATTAATCCACT 480
Qy 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCAATTAACCAAGTCCCTTAAAAAA 540
Db 481 CGAAAGCTATCATGTAAATGTTTAAAGAAACATCAATTAACCAAGTCCCTTAAAAAA 540
Qy 541 CAAGCATATTTTGGAAAGACAAATTTATGTTACAGTTTACAAATCTTAAGAGCGACAA 600
Db 541 CAAGCATATTTTGGAAAGACAAATTTATGTTACAGTTTACAAATCTTAAGAGCGACAA 600
Qy 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGTTTCTTTTCAATTTCTGTAATTTGTT 660
Db 601 TTATATCGAAAGGTAAGCTATGACGTTCAAGTTTCTTTTCAATTTCTGTAATTTGTT 660
Qy 661 ATTGTTTTTATATATATTTCTTCTCTTACAAATAGAGTATTTCTCGAATTTTAAAA 720
Db 661 ATTGTTTTTATATATATTTCTTCTCTTACAAATAGAGTATTTCTCGAATTTTAAAA 720
Qy 721 ATGACTATAAAGTCAATTTTATATATAGAGACGAGTCTGTAATTTCTCGTCAAAAAATC 780
Db 721 ATGACTATAAAGTCAATTTTATATATAGAGACGAGTCTGTAATTTCTCGTCAAAAAATC 780
Qy 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840
Db 781 TTTCTGATTTTTTAAAGCTAGTTTGGCAACCTGTTTCTTTCAAGAAATTTGATTTT 840
Qy 841 TTCAAAAAAATAGTTATTTTCTCTTAAATTAAGAAACATTAAGAAATTTAGT 900
Db 841 TTCAAAAAAATAGTTATTTTCTCTTAAATTAAGAAACATTAAGAAATTTAGT 900
Qy 901 TGCAGACTAGCCCTAGATGTTTTCCCAATTAATTAATCAATCACTGTAATTTATTTG 960
Db 901 TGCAGACTAGCCCTAGATGTTTTCCCAATTAATTAATCAATCACTGTAATTTATTTG 960
Qy 961 GCCAGCCCCCAATTAATTTAAACCGAACTGAATGAGCGAAACCAATTTGAGCTAT 1020
Db 961 GCCAGCCCCCAATTAATTTAAACCGAACTGAATGAGCGAAACCAATTTGAGCTAT 1020
Qy 1021 TTCTCTAGATTTAGTAAAGGAGAGAGAGAGAAATCAAGTTTAACTCAATTTGCC 1080
Db 1021 TTCTCTAGATTTAGTAAAGGAGAGAGAGAGAAATCAAGTTTAACTCAATTTGCC 1080
Qy 1081 TGAGATGTGCGTTTGGCAACGATAGCCACCGTAATCATAGTCTAGTCTAGCTCA 1140
Db 1081 TGAGATGTGCGTTTGGCAACGATAGCCACCGTAATCATAGTCTAGTCTAGCTCA 1140
Qy 1141 GGTTCGGCACTCTCGTGTCACTCAGATGCACTACTACTGCTTCAACCGTTGCTC 1200
Db 1141 GGTTCGGCACTCTCGTGTCACTCAGATGCACTACTACTGCTTCAACCGTTGCTC 1200
Qy 1201 TTGTTTCATGCTCGAAGCCTTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Db 1201 TTGTTTCATGCTCGAAGCCTTGCTTATTTGAAACCAAGAGATACCTACCCAAACAT 1260
Qy 1261 CCATCTTACTCATGCACTTTCATGCAACGACATATGTTTCTCGAAGCAGATCAT 1320
Db 1261 CCATCTTACTCATGCACTTTCATGCAACGACATATGTTTCTCGAAGCAGATCAT 1320
Qy 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Db 1321 AAAGATCAACAAGCTAGCTTCTCCGCTAGCTTCCCTCTCTCTGCGCATCTTTT 1380
Qy 1381 CGTCACACCATG 1394
Db 1381 CGTCACACCATG 1394

Db 1381 CGTCACACCATG 1394
RESULT 3
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9
Query Match 17.1%; Score 238.8; DB 20; Length 255;
Best Local Similarity 97.2%; Pred. No. 3.2e-42;
Matches 243; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1145 CGGAGCTCTGCTGATCTCAATGCACTACTACTAGCTTGTCAACCGTTGCTGT 1204
Db 5 CCGGAGTCCGGTGTATCTCAATGCACTACTACTAGCTTGTCAACCGTTGCTGT 64
Qy 1205 TCCATGTCGAAGCTTGGCTTATTTGAAACCAAGAGATTAATCTTCCAAACATTCAT 1264
Db 65 TCCATGTCGAAGCTTGGCTTATTTGAAACCAAGAGATTAATCTTCCAAACATTCAT 124
Qy 1265 CTTACTCATGCACTTCCATGCAACGACACATATGTTTCTGAAACCAATTCATTAAG 1324
Db 125 CTTACTCATGCACTTCCATGCAACGACACATATGTTTCTGAAACCAATTCATTAAG 184
Qy 1325 ATCAACAAGCTAGCGTTTCCGCTAGCTTCCCTCTCTCTCTGCGCATCTTTTCTGTC 1384
Db 185 ATCAACAAGCTAGCGTTTCCGCTAGCTTCCCTCTCTCTCTGCGCATCTTTTCTGTC 244
Qy 1385 CACCAACATG 1394
Db 245 CACCAACATG 254
RESULT 4
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3


```
/ LENGTH: 158
/ TYPE: DNA
/ ORGANISM: Zea mays
US-10-713-381-3

Query Match
Best Local Similarity 10.5%; Score 146; DB 20; Length 158;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1155 CGTGTCTTCACATGGCATACATGCTGTGTCACCGCTTGTC-TTGTTCACATGTC 1213
DB 1 CGTGTCTTCACATGGCATACATGCTGTGTCACCGCTGTGTCACATGTC 60

QY 1214 CAAGCCTTGCTTATTTGGAACCAAGAGATACCTACCCCAACAATTCATCTTACTCAT 1273
DB 61 CAAGCCTTGCTTATTTGGAACCAAGAGATACCTACCCCAACAATTCATCTTACTCAT 120

QY 1274 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 1311
DB 121 GCAACTTCATGCAACACGACATATGTTTCTGTAAC 158

RESULT 5
US-10-425-115-134230
/ Sequence 134230, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21 (53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 134230
/ LENGTH: 1261
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_53903C.1
US-10-425-115-134230

Query Match
Best Local Similarity 5.8%; Score 80.8; DB 20; Length 1261;
Matches 97; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 792 TTTAAGAGCTAGTTTGGAACCCCTGTTCTTTCAGAAATTTGATTTTCAAAAAA 851
DB 999 TCTTAAGGCTAGTTTGGAACCTTATTTTAAAGATTTTCTTTTAAAGATAA 1058

QY 852 TTAGTTATTTTCTCTTAAATAAGAAAACACTTAGAAAATAGAGTTGCCAGACTAG 911
DB 1059 TTAGTTATTTTCTCTTAAATAAGAAAATCTTGGAAAATTTGGAGTTGCTAAACTAG 1118

QY 912 CCCT 915
DB 1119 CCCT 1122

RESULT 6
US-10-739-930-1857
/ Sequence 1857, Application US/10739930
/ Publication No. US20040216190A1
/ GENERAL INFORMATION:
/ APPLICANT: Kovalic, David K.
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
/ FILE REFERENCE: 38-21 (53377)B
/ CURRENT APPLICATION NUMBER: US/10/739,930
/ CURRENT FILING DATE: 2003-12-18
```

```
/ NUMBER OF SEQ ID NOS: 11088
/ SEQ ID NO 1857
/ LENGTH: 1663
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: ZEAMA-23APR03-CLUSTER1429_8
US-10-739-930-1857

Query Match
Best Local Similarity 5.7%; Score 78.8; DB 20; Length 1663;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1313 AATCCATTAAAGATCAACACAGCTAGCGTTCTCCGCTTCCCTCTCTCTGCGG 1372
DB 1 AATCCATTAAAGATCAACACAGCTAGCGTTCTCCGCTTCCCTCTCTCTGCGG 60

QY 1373 ATCTTTTGTCTCACCACCATG 1394
DB 61 ATCTTTTGTCTCACCACCATG 82

RESULT 7
US-10-425-114-23340
/ Sequence 23340, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaska, Jack E
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21 (53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 23340
/ LENGTH: 320
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB3595-039-G8_FLI
US-10-425-114-23340

Query Match
Best Local Similarity 5.3%; Score 74; DB 18; Length 320;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 791 TTTAAGAGCTAGTTTGGAACCCCTGTTCTTCAAAAGATTTGATTTTCAAAAAA 850
DB 62 TTTTAAGGCTAGTTTGGAATCTCATTTTTCAGAGATTTTATTTTCTTAAGAA 121

QY 851 ATTAGTTATTTTCTCTTAAATAAGAAAACACTTAGAAAATAGAGTTGCCAGACTA 910
DB 122 ATTAGTTATTTTCTCTTGAAGAAAATAGAAATCCCTTGGAAATTAGAGTTCTAAACA 181

QY 911 GCCCTAGAA 920
DB 182 GCCCTAAAT 191

RESULT 8
US-10-425-114-16264
/ Sequence 16264, Application US/10425114
/ Publication No. US2004003488A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
```



```

; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16264
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3062-023-G10_FLI
US-10-425-114-16264
```

```
Query Match          5.3%; Score 74; DB 18; Length 624;
Best Local Similarity 73.1%; Pred. No. 5.8e-06;
Matches 95; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 791 TTTTAAGAGCTAGTTGGCAACCCCTGTTCTTCAAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 374 TTTTGGGCTAGTTGGAATCTCATTTTCAAGAGATTTTATTTCTTAAGAA 433
    |||||
QY 851 ATTAGTTATTTCTTTTAAATAGAAACCTTAGAAAAATGAGTTGCCAGACTA 910
    |||||
DB 434 ATTAGTTATTTCTTTGAGAAATGAAATCCCTTGGAAATGAGTTCTAAACGA 493
    |||||
QY 911 GCCCTAGAAAT 920
    |||||
DB 494 GCCCTTAATT 503
    |||||
```

```

RESULT 9
US-10-425-115-141826
; Sequence 141826, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 141826
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_60829C.1
US-10-425-115-141826
```

```
Query Match          5.2%; Score 73; DB 20; Length 1326;
Best Local Similarity 69.1%; Pred. No. 1.4e-05;
Matches 114; Conservative 0; Mismatches 50; Indels 1; Gaps 1;
```

```

QY 756 TGTCGTAGATTCCTTCAGAAAAATCTTGTGATTTTGAAGCTAGTTGGCAACCT 815
    |||||
DB 1090 TGAAGTGGGTGGTGGCTGCTGATTTCACTTTAACTGTAAGACTTAATTTGGAAACCA 1149
    |||||
QY 816 GTTCTTTCAAGAAATTTGATTTTCAAAAAATGATTTATTTCTTTATATAA 875
    |||||
DB 1150 ATTT-ATTCACCTGATTTTCTTCTTCTAGGAAATATGTTCTATTTCCCTTGAGAAAA 1208
    |||||
QY 876 TAGAAACACTTAGAAAAATAGAGTTCAGACTAGCCCTAGAAAT 920
    |||||
DB 1209 TAGGAATCCCTTGAAGAAAAATAGTTTCCAACTAGCCCTTAAT 1253
    |||||
```

```

RESULT 10
US-10-425-114-32493
; Sequence 32493, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 32493
; LENGTH: 2445
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73402B09_FLI
US-10-425-114-32493
```

```
Query Match          5.2%; Score 72.8; DB 18; Length 2445;
Best Local Similarity 73.6%; Pred. No. 2.1e-05;
Matches 106; Conservative 0; Mismatches 37; Indels 1; Gaps 1;
```

```

QY 792 TTTAAGAGCTAGTTGGCAACCCCTGTTCTT-TCAGAAATTTGATTTTCAAAAAA 850
    |||||
DB 1623 TCTTAGGCTAGTTGAGAACCCCTTTTTCCTCAAGATTTTCATTTTCAAGAAAA 1682
    |||||
QY 851 ATTAGTTATTTCTTTTAAATAGAAACACTTAGAAAAATGAGTTGCCAGACTA 910
    |||||
DB 1683 ATTAGTTATTTCTTTCTTCTTCTAGGAAATGGAATCCCTTAAGAAAAATGAGTTTCAAACTA 1742
    |||||
QY 911 GCCCTAGAAATGTTTCCCAATATA 934
    |||||
DB 1743 GCCCTTAATGTTTTCATGATA 1766
    |||||
```

```

RESULT 11
US-10-425-115-83293/C
; Sequence 83293, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 83293
; LENGTH: 2729
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2729)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_175978C.1
US-10-425-115-83293
```

```
Query Match          5.2%; Score 72.8; DB 20; Length 2729;
Best Local Similarity 60.7%; Pred. No. 2.2e-05;
Matches 136; Conservative 0; Mismatches 87; Indels 1; Gaps 1;
```


This Page Blank (unfiled)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 4347.69 Seconds
(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1

Perfect score: 1394
Sequence: 1 ccctggcgtcctcctcgaataa.....cttttcgtccaccaccatcg 1394

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

ESR:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	952	68.3	963	9	CC656933 OGMQ20TM
2	760.4	54.5	915	9	CG224225 OGIAG08TV
3	501.2	36.0	687	9	CC656933 OGMQ20TV
4	96	6.9	715	9	CG252571 OG4B805TC
5	92	6.6	967	9	CL235046 ZMMBB057
6	88	6.3	814	9	CG048704 PU1019TB
7	87.8	6.3	754	9	CG414922 ZMMBB029
8	86.8	6.2	950	8	CC439901 PUHRI15TB
9	86.6	6.2	652	8	CC384247 PUHOC67TB
10	86.6	6.2	797	8	CC400575 PUHLU61TD
11	85.6	6.2	820	8	CC400574 PUHLU61TB
12	85.4	6.1	471	9	CG103452 PUHBI9TB
13	85.2	6.1	765	9	CG082135 PUFOX12TD
14	85.2	6.1	781	9	CG630219 OGUCG53TV
15	85.2	6.1	815	9	CG349565 OG0F183TV
16	85.2	6.1	834	9	CG349565 OG0F183TV
17	85	6.1	789	8	CC433618 PUHHP17TD
18	84.6	6.1	1092	8	CNS020K7
19	84	6.0	793	8	BZ816381
20	84	6.0	1078	9	CL997678
21	83.8	6.0	530	9	CG201774
22	83.6	6.0	722	9	CG333914
23	83.6	6.0	722	9	CG333929
24	83.6	6.0	781	9	CG034985

25	83.6	6.0	861	8	BZ797976
26	83.6	6.0	865	8	CC430754
27	81.8	5.9	861	9	CG102092
28	81.8	5.9	947	8	CC435780
29	81.8	5.9	981	8	BZ784278
30	81.8	5.9	1017	9	CC620594
31	81.6	5.9	649	9	CC613918
32	81.6	5.9	733	8	BZ778636
33	81.6	5.9	925	9	CG071791
34	81.6	5.9	960	8	BZ676889
35	81.6	5.9	994	8	CC003943
36	81.6	5.9	1016	9	CL996481
37	81.6	5.9	1022	9	CL984151
38	81.4	5.8	999	8	CC385762
39	81.2	5.8	756	9	CG221653
40	81	5.8	1101	9	CNS001T2
41	80.6	5.8	699	8	BZ969930
42	80.6	5.8	832	9	CC676387
43	80.6	5.8	839	9	CG246849
44	80.6	5.8	843	9	CG254365
45	80.6	5.8	852	8	CC385881

ALIGNMENTS

RESULT 1
LOCUS CC656933 963 bp DNA linear GSS 19-JUN-2003
DEFINITION OGMQ20TM ZM 0.7 1.5 KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.

ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Buddiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
1 (bases 1 to 963)

TITLE
JOURNAL
COMMENT
Unpublished (2002)
Other GSSs: OGMQ20TV
Contact: Cathy Whitelaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source
1..963
Location/Qualifiers

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0554D15"
/note="Vector: pBCK-; Site 1: HincII, 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 68.3% Score 952; DB 9; Length 963;
Best Local Similarity 99.9%; Pred. No. 3.8e-191;

Matches 963; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

404 ACCAGCATGTCTAGTCTCACTATTTGGCACTACCAAGCAAGATTTAAAAAATACCC 463
1 ACCAGCATGTCTAGTCTCACTATTTGGCACTACCAAGCAAGATTTAAAAAATACCC 60


```
QY 464 AAAGTAACATCACTCGAAGGCTATCATGTAATGTTTAAAGAAACATCTATTAACC 523
DB 61 AAAGTAACATCACTCGAAGGCTATCATGTAATGTTTAAAGAAACATCTATTAACC 120
QY 524 AGCATCCTCTTAAACAAAGCATATTTCCGAAAGACAAATATGTTACAGTTTACAA 593
DB 121 AGCATCCTCTTAAACAAAGCATATTTCCGAAAGACAAATATGTTACAGTTTACAA 180
QY 584 CATCTAAGACGCGCAAAATATATCGAAAGGTAGCTATGACGTTTCAGATTTTCTTTTC 643
DB 181 CATCTAAGACGCGCAAAATATATCGAAAGGTAGCTATGACGTTTCAGATTTTCTTTTC 240
QY 644 ATCTCTGTAATTTGTTATGTTTATATATACATTTTCTCTCTCAATAGAGTATTT 703
DB 241 ATCTCTGTAATTTGTTATGTTTATATATACATTTTCTCTCTCAATAGAGTATTT 300
QY 704 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATATTAAGACAGCATGTCGTAG 763
DB 301 TCTTCGATTTTAAATGACTATTAAGTCAATTTTATATTAAGACAGCATGTCGTAG 360
QY 764 ATCTCGTTCAAAATCTTCTGATTTTAAAGAGTATGCGAACCTGTTCTTT 823
DB 361 ATCTCGTTCAAAATCTTCTGATTTTAAAGAGTATGCGAACCTGTTCTTT 420
QY 824 CAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATATTAATAGAAAAC 883
DB 421 CAAGAATTTGATTTTCAAAAAAATAGTTTATTTCTCTTATATTAATAGAAAAC 480
QY 884 ACTTGAAGAAATAGAGTGGCAGAGTACCCCTAGAAATGTTTCCCAATTAATCAATCA 943
DB 481 ACTTGAAGAAATAGAGTGGCAGAGTACCCCTAGAAATGTTTCCCAATTAATCAATCA 540
QY 944 CTGTGTATTAATTTTGGCCAGCCCATTAATTTTAAACCGAACTGAAATCGAGCA 1003
DB 541 CTGTGTATTAATTTTGGCCAGCCCATTAATTTTAAACCGAACTGAAATCGAGCA 600
QY 1004 AACCAATCTGAGCTATTTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAAATCAAG 1063
DB 601 AACCAATCTGAGCTATTTCTCTAGATTTAGTAAAGAGAGAGAGAGAGAAATCAAG 660
QY 1064 TTTTAAGCTATGTCCTCGAGATGTCGGTTTGGCAAGATGCAACCGTATCATAGCT 1123
DB 661 TTTTAAGCTATGTCCTCGAGATGTCGGTTTGGCAAGATGCAACCGTATCATAGCT 720
QY 1124 CAAAGTGCCTAGTCAAGTTCGGGAGCTTCGTGTATCTCAATGCAATGCTACTATGTC 1183
DB 721 CAAAGTGCCTAGTCAAGTTCGGGAGCTTCGTGTATCTCAATGCAATGCTACTATGTC 780
QY 1184 TTGTTCAACGCTGCTGTTGTCATGTCGAAGCTTGCGTATTTCTGAACCAAGAGAT 1243
DB 781 TTGTTCAACGCTGCTGTTGTCATGTCGAAGCTTGCGTATTTCTGAACCAAGAGAT 840
QY 1244 ACCTACTCCCAACCAATTCATCTTACTCATGCAACTTCATGCAACACGCAATATGTT 1303
DB 841 ACCTACTCCCAACCAATTCATCTTACTCATGCAACTTCATGCAACACGCAATATGTT 900
QY 1304 TCTTGAACCAATTCATTAAGATCAACAGCTAAGGTTCTCCGCTACTTCCCTCTCT 1363
DB 901 TCTTGAACCAATTCATTAAGATCAACAGCTAAGGTTCTCCGCTACTTCCCTCTCT 959
QY 1364 CCTC 1367
DB 960 CCTC 963
```

```
RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
LOCUS CG224225
DEFINITION OGIAG08YV_ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0716B15,
GENOMIC CG224225 genomic survey sequence.
ACCESSION CG224225.1 GI:34124113
```

```
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 915)
AUTHORS Whitehead, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,
Reinick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nundberg, A., Rodolins, D. and Lakey, N.
CONSORTIUM Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Other GSSs: OGIAG08TH
Contact: Cathy Whitehead
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitehead@cigr.org
Seq primer: TF
Classes: sheared ends.
FEATURES
source
location/Qualifiers
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0716B15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HindIII, 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 54.5%; Score 760.4; DB 9; Length 915;
Best Local Similarity 99.7%; Pred. No. 1.5e-150;
Matches 772; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 621 TGACGTCAGATTTTCTTTTCAATTTGTTATTTGTTATTTTATATACATTTT 680
DB 1 TGACGTCAGATTTTCTTTTTC-TTCTTGTTATTTGTTATTTTATATACATTTT 59
QY 681 CTTCCTTCAATAGATGATTTTCTCGATTTTATTAATGACTATTAAGTCAATTTT 740
DB 60 CTTCCTTCAATAGATGATTTTCTCGATTTTATTAATGACTATTAAGTCAATTTT 119
QY 741 ATATAGAGCAGCAGTGTGATGATTTCTGTCGTTCAAAAATCTTTCGATTTTAAAGC 800
DB 120 ATATAGAGCAGCAGTGTGATGATTTCTGTCGTTCAAAAATCTTTCGATTTTAAAGC 179
QY 801 TAGTTGGCAACCCGTTTCTTCAAGAAATTTGATTTTCAAAAAAATTAGTTTAT 860
DB 180 TAGTTGGCAACCCGTTTCTTCAAGAAATTTGATTTTCAAAAAAATTAGTTTAT 239
QY 861 TTTCTCTTATTAATTAAGAAACACTTAGAAATATAGATTGCGAGCTAGCCCTAGAT 920
DB 240 TTTCTCTTATTAATTAAGAAACACTTAGAAATATAGATTGCGAGCTAGCCCTAGAT 299
QY 921 GTTTTCCCAATTAATCAATCAGTGTATTAATTTTGGCAGCCCATTAATTTT 980
DB 300 GTTTTCCCAATTAATCAATCAGTGTATTAATTTTGGCAGCCCATTAATTTT 359
QY 981 AAACGAATGAATTCAGAGCAACCAATCTGAGTATTTCTAGATTAGTAAAG 1040
DB 360 AAACGAATGAATTCAGAGCAACCAATCTGAGTATTTCTAGATTAGTAAAG 419
QY 1041 GGAGAGAGAGAGAAATCAATTAAGTATGTCCTGAGATGTCGGTTTGGCA 1100
DB 420 GGAGAGAGAGAGAAATCAATTAAGTATGTCCTGAGATGTCGGTTTGGCA 479
QY 1101 CGATAGCCACCGTATCATAGTCTATAGTGTCTTACGTCAGGTTGGCAGCTCTGTGTC 1160
DB 480 CGATAGCCACCGTATCATAGTCTATAGTGTCTTACGTCAGGTTGGCAGCTCTGTGTC 539
```



```

QY 1161 ATCTACATGGCATACTACATGCTTGTTCACCGCTTGCTGTGTCATGCTCCAGCCT 1220
|
|
|
Db 540 ATCTACATGGCATACTACATGCTTGTTCACCGCTTGCTGTGTCATGCTCCAGCCT 599
|
|
|
QY 1221 TGCCTATTCTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCATGCAACT 1280
|
|
|
Db 600 TGCTATTCTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCATGCAACT 659
|
|
|
QY 1281 CCATGCAACACGACATATGTTCTCTGAAACCAATTCATTAAGATACACACAGCTACGC 1340
|
|
|
Db 660 CCATGCAACACGACATATGTTCTCTGAAACCAATTCATTAAGATACACACAGCTACGC 719
|
|
|
QY 1341 TTCTCCCGTAGCTGCTCTCTCTGCGCATCTTTTCTGTCACACACATG 1394
|
|
|
Db 720 TTCTCCCGTAGCTGCTCTCTCTCTGCGCATCTTTTCTGTCACACACATG 773
|
|
|
RESULT 3
CC656939/c 687 bp DNA linear GSS 19-JUN-2003
LOCUS CC656939
DEFINITION OGMDQ20TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION CC656939
VERSION CC656939.1 GI:32060231
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 687)
Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
OTHER_GSS: OGMDQ20TM
COMMENT Contact: Cathy Whitefaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..687
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0554D15"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 36.0%; Score 501.2; DB 9; Length 687;
Best Local Similarity 99.4%; Pred. No. 1,2e-95;
Matches 503; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 889 GAAAAATAGAGTGGCCAGACTAGCCCTAGAGATGTTTCCCAATTAATTAATCAATCACTG 948
|
|
|
Db 687 GAAAAATAGAGTGGCCAGACTAGCCCTAGAGATGTTTCCCAATTAATTAATCAATCACTG 628
|
|
|
QY 949 TATAATTAATTTGGCCAGCCCATTAATTAATTAACGGAATCGAATCGAGGAACCA 1008
|
|
|
Db 627 TATAATTAATTTGGCCAGCCCATTAATTAATTAACGGAATCGAATCGAGGAACCA 568
|
|
|
QY 1009 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTT 1068
|
|
|
Db 567 AATCTGAGCTATTTCTCTAGATTAGTAAAAAGGAGAGAGAGAGAGAAATCAGTTT 508
|
|
|

```

```

QY 1069 AGTCATTGTCCTTGAGATGAGCGCTTTGGCAACGATAGCCATATAGCTATAG 1128
|
|
|
Db 507 AGTCATTGTCCTTGAGATGAGCGCTTTGGCAACGATAGCCATATAGCTATAG 448
|
|
|
QY 1129 GTGCTACGTCAGAGTTCGGAGCTCTGTCATCTCAATGAGCATACTAGCTTGT 1188
|
|
|
Db 447 GTGCTACGTCAGAGTTCGGAGCTCTGTCATCTCAATGAGCATACTAGCTTGT 388
|
|
|
QY 1189 CAACGCTGCTTTGTTCCATGTCACAGCTTTCCTATTCTGAACCAAGAGATACCT 1248
|
|
|
Db 387 CAACGCTGCTTTGTTCCATGTCACAGCTTTCCTATTCTGAACCAAGAGATACCT 328
|
|
|
QY 1249 CTCCTCAACCAATTCATCTTATCATGCACTTCCTGCAACCAACGAGATATGTTCC 1308
|
|
|
Db 327 CTCCTCAACCAATTCATCTTATCATGCACTTCCTGCAACCAACGAGATATGTTCC 268
|
|
|
QY 1309 AACCAATTCATTAAGATACAAACAGCTAGCTTCCGCTAGCTTCCCTCTCTCT 1368
|
|
|
Db 267 AACCAATTCATTAAGATACAAACAGCTAGCTTCCGCTAGCTTCCCTCTCTCTCT 208
|
|
|
QY 1369 GCCGATCTTTTGTGTCACACCATG 1394
|
|
|
Db 207 GCCGATCTTTTGTGTCACACCATG 182
|
|
|
RESULT 4
CG252571/c 715 bp DNA linear GSS 25-AUG-2003
LOCUS CG252571
DEFINITION OG48B057C ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0809B10,
genomic survey sequence.
ACCESSION CG252571
VERSION CG252571.1 GI:34154661
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 715)
Whitefaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nuneberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
COMMENT Contact: Cathy Whitefaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitefaw@tigr.org
Seq primer: TF
Class: sheared ends.
FEATURES
source location/Qualifiers
1..715
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0809B10"
/clone_1lb="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"
ORIGIN
Query Match 6.9%; Score 96; DB 9; Length 715;
Best Local Similarity 80.1%; Pred. No. 8.4e-10;
Matches 125; Conservative 0; Mismatches 30; Indels 1; Gaps 1;
QY 135 TCTCCACCTCGGCTGACATTTCTTGATGTGGTGTCCATCTGACGAGGCCCA 194
|
|
|
Db 356 TCACACACCGCTCTTGCTCTTTCTTGATGTGGGCGGCTCCCGCTCATCGAGGCCCA 297
|
|
|
QY 195 TCACACACCTTTTGGGACACCATCAAGGCCCTTTGATGTGGCCCAAGAGAGCTATCGG 254
|
|
|

```



```

Db      296 TAAGGACACCTGCGAGAACCCATTAAAGCCTTTCGATGCGCCAGGACAT-CGGGG 238
Qy      255 TCCTGGATCCAGGAGATATATGTCCTCCCAATC 290
Db      237 TCGTGTGACCCGGGGATATATATCCCAACC 202

RESULT 5
LOCUS   CL235046
DEFINITION ZM5001r ZM5001 (HindIII) Zea mays genomic clone
ACCESSION CL235046
VERSION   CL235046.1 GI:40891729
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
AUTHORS   Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE     Sequencing of the maize genome at PGR (2003c)
JOURNAL   Unpublished (2003)
COMMENT   Contact: Bharti,A.K.
           Dr.Joachim Messing's lab
           The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
           University
           190 Frelinghuysen Road, Piscataway, NJ 08854, USA
           Tel: 732 445 3801
           Fax: 732 445 5735
           Email: bharti@waksman.rutgers.edu
           Seq primer: SP6
           Class: BAC ends
           High quality sequence start: 64.
           Location/Qualifiers
             1. 967
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /culivar="B73"
             /db_xref="taxon:4577"
             /clone="ZM5001"
             /lab_host="E. coli DH10B"
             /clone_lib="ZM5001 (HindIII)"
             /note="Vector: pUC19; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match      6.6%; Score 92; DB 9; Length 967;
Best Local Similarity 78.6%; Pred. No. 6e-09;
Matches 110; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy      150 TGCACATTTTTCGATGTCGGTTCCTGACCGAGCCCATCAGACACCTTTGG 209
Db      827 TCGCCCTTCTTGATACCGCGCGTTCCTGATGAGGCCCATCCACCTTTTGG 886
Qy      210 GACACCATCAAGGGCTTCGATGCGCCAGAGAGATGCGGTGCGTGATCCAGG 269
Db      887 GTGGCCCATCAATGCTTTCGATGTCACCGGAGATGATCCGGGCGCTGACCTGGG 946
Qy      270 GGATATATGTCCTCCCAAT 289
Db      947 GGATATCTGTCCCAATAGT 966

RESULT 6
LOCUS   CG048704
DEFINITION 814 bp DNA linear GSS 19-AUG-2003
ACCESSION CG048704
VERSION   CG048704.1 GI:33920884

```

```

KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE Whiteley,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
AUTHORS   Renick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
           Benetren,J.
TITLE     Maize Genomics Consortium
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: PU1019TD
           Contact: Cathy Whiteley
           TIGR
           9712 Medical Center Drive, Rockville, MD 20850, USA
           Tel: 301-838-5843
           Fax: 301-838-0208
           Email: whiteley@tigr.org
           Seq primer: TR
           Class: sheared ends.
           Location/Qualifiers
             1. 814
             /organism="Zea mays"
             /mol_type="genomic DNA"
             /strain="B73"
             /db_xref="taxon:4577"
             /clone="ZM5001"
             /clone_lib="ZM_0_6_1_0_KB"
             /note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
             Cot selected genomic DNA library"

ORIGIN
Query Match      6.3%; Score 88; DB 9; Length 814;
Best Local Similarity 62.7%; Pred. No. 4.2e-08;
Matches 153; Conservative 0; Mismatches 90; Indels 1; Gaps 1;

Qy      713 TTTATTAATGACTTTAAAGCATTTTATATTAAGACAGCGATGCTGATTCGTT 772
Db      487 TTTAAGACAGCTTGATGATGAGAAAGATATAGAAATATCTTTAGGAATGTTA 546
Qy      773 CAAAAATCTTCTGATTTTAAAGAGCTAGTTGGCAACCTGTCTTCAAGAAAT 832
Db      547 TAAAGACAGAAACATTTCTTAAGAGCTAGTTGATCCATTTTCAAGAAAT 606
Qy      833 TTGATTTTTCAAA-AAAATAGTTTATTTCTTTATTAATAGAAACATTAGA 891
Db      607 TTCATTTCAAGGAGAAATATGTTATTTCTTTGTAATAGTAATCAGCGGA 666
Qy      892 AAATAGAGTTCAGACTAGCCCTAGAAATGTTTCCCAATTAATCAATCACTGTAT 951
Db      667 AAATAGAGTTCAGAACTAGCCCTAAGGTATATATCCGATATCCCTATCTCT 726
Qy      952 AATT 955
Db      727 ATTT 730

RESULT 7
LOCUS   CG414922/c
DEFINITION 754 bp DNA linear GSS 08-SEP-2003
ACCESSION CG414922
VERSION   CG414922.1 GI:34505144
KEYWORDS  GSS.
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE Bharti,A.K., Soderlund,C., Soderlund,C., Bharti,A.K., Messing,J.
AUTHORS   Yu,Y., Kim,H.R., Halfeld,J.,
           1 (bases 1 to 754)

```


Matches	131:	Conservative	0:	Mismatches	74:	Indels	0:	Gaps	0:
Qy	714	TTATTAATAATGACTATAAAGTCATTTTATATATAGACGACGATGCTAGATTCCTGCT	773						
Db	341	TAAATTAATAGTAGTACTAGTACATCTTAAATCGACATATATATGATTAATATATGCTA	400						
Qy	774	AAAAATCTTCGATTTTTTTTAAGAGCTAGTTGGCAACCCGTTCTCTTCAAGAATTT	833						
Db	401	AGATCTTATTTATTTATATATGTTAAGAGCTAGTTGGCAACCTATTTTTCTAAGAGATTTT	460						
Qy	834	TGATTTTTTCAAAAAAATATAGTTATTTCTCTTATATAATAGAAAAACCTAGAAAA	893						
Db	461	TATTTTTTCAAGTGAATTTAGTTATTTTCCATGAGAAAAATGAATTAATTTGGAAAA	520						
Qy	894	ATGAGTTCCAGACTAGCCCTAGA	918						
Db	521	ATTAGTTTCCAAACTAGCCCTAAA	545						
RESULT 10									
CC400575/c									
LOCUS									
DEFINITION	CC400575	797 bp	DNA	linear	GSS 19-MAY-2003				
	PURLUGITD ZM 0.6_1.0 KB	Zea mays	genomic	clone ZMMB7a480L01,					
ACCESSION	CC400575								
VERSION	CC400575.1	GI:30880665							
KEYWORDS	GSS,								
SOURCE	Zea mays								
ORGANISM	Zea mays								
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD								
	clade; Panicoideae; Andropogoneae; Zea.								
REFERENCE	1 (bases 1 to 797)								
AUTHORS	Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,								
	Benick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and								
	Bennerzen,J.								
TITLE	Maize Genomics Consortium								
JOURNAL	Unpublished (2003)								
COMMENT	Other_GSSs: PURLUGITB								
	Contact: Cathy Whitelaw								
	TIGR								
	9712 Medical Center Drive, Rockville, MD 20850, USA								
	Tel: 301-838-5843								
	Fax: 301-838-0208								
	Email: whitelaw@tigr.org								
	Seq primer: TP								
	Class: sheared ends.								
FEATURES									
source	location/Qualifiers								
	1..797								
	/organism="Zea mays"								
	/mol_type="genomic DNA"								
	/strain="B73"								
	/db_xref="taxon:4577"								
	/clone="ZMMB7a480L01"								
	/clone_lib="ZM 0.6_1.0_KB"								
	/note="Vector: PCR4-TOPO, Site_1: EcoRI, 0.6-1.0 kb high								
	COT selected genomic DNA library"								
ORIGIN									
Query Match	6.2%	Score 86.6:	DB 9:	Length 797:					
Beet Local Similarity	59.7%:	Pred. No. 8.3e-08:							
Matches 163:	Conservative	0:	Mismatches 109:	Indels 1:	Gaps 1:				
Qy	713	TTTATTAATAATGACTATAAAGTCATTTTATATTAAGAGCAGCATGCTAGATTCCTGCT	772						
Db	577	TTTAGAGGACGTTGGTAGAGATGAGAAAGATATAGAGAAATATATCTTTAGAGATGTTA	518						
Qy	773	CAAAATCTTCTGATTTTTTTTAAGAGCTAGTTGGCAACCCCTGTTCTTTCAAGAATT	832						
Db	517	TAAAGGACAGAAACATTTTTTTTAAGAGCTAGTTGGTAATCCATTTTTTTTCAAGGATT	458						
Qy	833	TTGATTTTTTCAAAA-AAATTAGTTTATTTCTCTTATATAAATAGAAAACTTGGAA	891						

Db	457	TTCAATTCTCACAAGGAAAAAATAGTTTATTTCCTTTTGGTAAATGGTATCACTCGGA	338
Qy	892	AAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTTCCCAATAATTAACAATCACTGTGTAT	951
Db	397	AAATAGAGTTCTCAAACTAGCCCTTAAGGTTATATCCGATTAATCCCTATCTCATTTCTCT	338
Qy	952	AATATTTGGCCAGCCCATTAATTAATTTAAAC	984
Db	337	ATTTCAAATTCACATTGGCAAAATAGTGCATC	305
RESULT 11			
LOCUS	CC400574	820 bp	DNA linear GSS 19-MAY-2003
DEFINITION	PHLU61TB ZM 0.6_1.0 KB Zea mays genomic clone ZMBRta480L01,		
ACCESSION	CC400574		
VERSION	CC400574.1		
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Zea mays		
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 820)		
AUTHORS	WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Utecherback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennezeren,J.		
TITLE	Maize Genomics Consortium		
JOURNAL	Unpublished (2003)		
COMMENT	Other_GSSs: PHLU61TD		
	Contact: Cathy Whitelaw		
	TIGR		
	9712 Medical Center Drive, Rockville, MD 20850, USA		
	Tel: 301-838-5843		
	Fax: 301-838-0208		
	Email: whitelaw@cigr.org		
	Seq primer: TR		
	Class: sheared ends.		
FEATURES			
source	location/Qualifiers		
	1..820		
	/organism="Zea mays"		
	/mol_type="genomic DNA"		
	/strain="B73"		
	/db_xref="taxon:45777"		
	/clone_lib="ZMBRta480L01"		
	/clone_1lb="ZM 0.6_1.0 KB"		
	/note="Vector: pCR4-T0B0; Site_1: EcoRI; 0.6-1.0 kb high		
	Cot selected genomic DNA library"		
ORIGIN			
	Query Match	6.2%;	Score 86.6; DB 8; Length 820;
	Best Local Similarity	59.7%;	Pred. No. 8.3e-08;
	Matches 163; Conservative	0;	Mismatches 109; Indels 1; Gaps 1;
Qy	713	TTTATAAATGACTATTAAGTCAATTTTATATAAAGACAGCATGTCGATGTCGTT	772
Db	374	TTTAGAGAGCGTTGGTAGAGATGGAAGAATATAGAGAAAATATCTTTTAGAATGTTA	433
Qy	773	CAAAAATCTTCGATTTTTTTTAAAGCTAGTTGGCAACCTGTTCTTCAAGAATT	832
Db	434	TAAAGACAAGAAACAATTTTAAAGCTAGTTGGTAATCCCATTTTTTCAAGGATT	493
Qy	833	TTGATTTTTTCAAAA-AAAATGAGTTTATTTCTCTTATAAATGAAAACTTAGAA	891
Db	494	TTCAATTCTCACAAGGAAAAATATGTTTATTTCCCTTTTGTAAATGATATCTCGGA	553
Qy	892	AAATAGAGTTGCCAGACTAGCCCTAGAAATGTTTTCCCAATAATTAACAATCACTGTGTAT	951
Db	554	AAATAGAGTTCTCAAACTAGCCCTTAAGGTTATATCCGATTAATCCCTATCTCATTTCTCT	613
Qy	952	AATATTTGGCCAGCCCATTAATTAATTTAAAC	984

Db	614	ATTCAATTCACCTTGCAGAAATGATGCATC	646
RESULT 12	CG103452	471 bp	DNA
LOCUS	PUBB19TB.ZM.0.6.1.0_KB_Zea_mays	genomic clone ZMMBTa0628D13,	GSS 20-AUG-2003
DEFINITION	genomic survey sequence.		
ACCESSION	CG103452		
VERSION	CG103452.1	GI:33985746	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 471)		
AUTHORS	Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T., Renick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Benner, J.		
TITLE	Maize Genomes Consortium		
JOURNAL	Unpublished (2003)		
COMMENT	Other GSSs: PUBB19TD Contact: Cathy Whitelaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@cigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers 1. 471 /organism="Zea mays" /mol_type="genomic DNA" /strain="873" /db_xref="taxon:4577" /clone="ZMMBTa0628D13" /clone_11b="ZM.0.6.1.0_KB" /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high cot selected genomic DNA library"		
FEATURES	Source		
ORIGIN	Query Match 6.1%; Score 85.4; DB 9; Length 471; Best Local Similarity 69.5%; Pred. No. 1.5e-07; Matches 116; Conservative 0; Mismatches 51; Indels 0; Gaps 0;		
Qy	789 TTTTAAAGCTAGTTGGCAACCGTCTTTCAAGAAATTTGATTTTTCAAAAA	848	
Db	216 TTTCTTAGACCTAGTTGGAAATCCATTTTTTTCAGAGATTACATTTTTTTCAGG	275	
Qy	849 AAATAGTTATTTCTCTTTATAAATAGAAAAACACTTAGAAAAATAGATGGCAGAC	908	
Db	276 AAATAGTTATTTCTCTTTGAAAAAATAGAAATACATGGGAAATAGTGTTCACAAAC	335	
Qy	909 TAGCCCTAGAAATGTTTCCCAATAAATTAACATCATCTGTATTAATT	955	
Db	336 TAGCGCTTAGACTATTTTCTAAAAAAGTAGATCCATGATTAAGATT	382	
RESULT 13	CG082135	765 bp	DNA
LOCUS	PUBOX12TD.ZM.0.6.1.0_KB_Zea_mays	genomic clone ZMMBTa0712B24,	GSS 20-AUG-2003
DEFINITION	genomic survey sequence.		
ACCESSION	CG082135		
VERSION	CG082135.1	GI:33964429	
KEYWORDS	GSS.		
SOURCE	Zea mays		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoidae; Andropogoneae; Zea.		

REFERENCE	1 (bases 1 to 765)
WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Renick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Benneken, J.	
TITLE	Maize Genomics Consortium
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: PUPFOX12B Contact: Cathy WhiteLaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@tigr.org Seq primer: TP Class: sheared ends.
FEATURES	
source	Location/Qualifiers 1..765 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMB7A0712B24" /clone_11b="ZM 0.6 1.0 KB" /note="Vector: PCR4-TOBO; Site 1: EcoRI; 0.6-1.0 kb high Cor selected genomic DNA library"
ORIGIN	
Query Match	6.1%; Score 85.2; DB 9; Length 765;
Best Local Similarity	71.8%; Pred. No. 1.6e-07;
Matches 125; Conservative	0; Mismatches 48; Indels 1; Gaps 1;
QY	787 ATTTTAAAGAGCTAGTTGGCAACCCGTTCTTTCGAAAGAAATTTGATTTTTCGAA 846
Db	255 ATCTTTATAGACTAGTTGGGAACCTTATTT-TCTCAGCGGATTTTATTTTTCAT 197
QY	847 AAAAAATAGTTTATTTCTCTTATATAAATAGAAAACACTTGAATAATAGATGGCAG 906
Db	196 GAAATATAGTTTATTTTCTCTTGAGAAAATATATCATCTTGAGAAAATATAGTTTCAA 137
QY	907 ACTAGCCCTGGAATGTTTTCCCATATAATATACATCACTGTGTATATTTATTTT 960
Db	136 ACTAGCCCTTAAGATTTTATATAAATAGAGAAAATATCGTCTATTTTTCATTG 83
RESULT 14	
CC630219/c	781 bp DNA linear GSS 19-JUN-2003
LOCUS	OCUCG531TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBKA0406J09,
DEFINITION	genomic survey sequence.
ACCESSION	CC630219
VERSION	CC630219.1 GI:32003354
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays
REFERENCE	WhiteLaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Benneken, J., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
AUTHORS	Maize Genomics Consortium for Maize Genomics Consortium (2002)
TITLE	Unpublished (2002)
JOURNAL	Other GSSs: OCUCG531TH
COMMENT	Contact: Cathy WhiteLaw TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whiteLaw@tigr.org Seq primer: TP Class: sheared ends. Location/Qualifiers

AX224394
LOCUS AX224394 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 1 from Patent WO0160997.
ACCESSION AX224394
VERSION AX224394.1 GI:15554636
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 60
DB 1155 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 1214

QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 3
AX224395
LOCUS AX224395 1394 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0160997.
ACCESSION AX224395
VERSION AX224395.1 GI:15554637
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES
source location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 60
DB 1155 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 1214

QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTGTAAC 1311

DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 4
BD062176
LOCUS BD062176 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062176
VERSION BD062176.1 GI:22607781
KEYWORDS JP 2001520523-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1998504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
PC C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 60
DB 1155 CGGTGATCTCACATGCGATACACTGCTGTTCAACCGTTGCTTGTTCATGTCG 1214

QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACACTCTCCCAACATCCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAACACGACATATGTTCTGTAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTGTAAC 1311

RESULT 5
BD062177
LOCUS BD062177 1394 bp DNA linear PAT 27-AUG-2002
DEFINITION Male tissue-preferred regulatory region and method of using same.
ACCESSION BD062177
VERSION BD062177.1 GI:22607782
KEYWORDS JP 2001520523-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 (bases 1 to 1394)

AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.

TITLE Male tissue-preferred regulatory region and method of using same

JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI-BRED INTERNATIONAL INC
PN JP 2001520523-A/2
PD 30-OCT-2001
PR 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAAAT, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.

FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 157; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCACTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 60
DB 1155 CGTGTATCTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 1214

QY 61 AAGCCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAAAACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCGAAC 1311

RESULT 6
AF360356 3343 bp DNA linear PLN 12-MAY-2001
LOCUS AF360356 Zea mays male fertility protein (Mg45) gene, complete cds.
DEFINITION
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Cloning of Mg45, a gene required for male fertility from Zea mays
unpublished
2 (bases 1 to 3343)
Fox, T.W., Trimmell, M.R. and Albertsen, M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
Location/Qualifiers
source 1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Mg45"
Join(<1392..1768,1898..2182,2280..2447,2534..>2942)

CDS
/gene="Mg45"
/product="male fertility protein"
Join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Mg45"
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKRLQWRGRDGIIVQYPHLFFAALALALVADPFGSLPLAEV
DYRPVGHLEAPYGEVSWPRDNASRLRGRLEFYGVGFPSIEFDLOGRYPAGLA
DGRVVRMGEBAEMETFAVWNPDMSEKAGVNSYTRQHEKEEFCGRPLGRFGE
TGEIVYADAVYGLMVGSGGVASVAREADDPIRFANDLVHNRGSYFEDTSMRY
SRKDHNLILBEGTRRLRYRPERSGVAVYAKGVFPNGVQISEDHQFLRSSTTC
RIMRWLBEGRAGEVEFANLPGFDPVNSNRGQVWAIIDCCTPDAQEVFARPLR
TLVFKPLSLKVLTWKARMTVLALIDEGRVVEVLDRGHEVWKIVSEVEVGRK
LWIGVAVHNHIVATIPYPLSD"

ORIGIN

Query Match 100.0%; Score 157; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 1.9e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCACTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 60
DB 1155 CGTGTATCTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 1214

QY 61 AAGCCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCATG 1274

QY 121 CAACCTTCATGCAAAACGACATATGTTCTCGAAC 157
DB 1275 CAACCTTCATGCAAAACGACATATGTTCTCGAAC 1311

RESULT 7
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS AX224396 Sequence 3 from Patent WO0160997.
DEFINITION
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G. and Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
source 1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 93.0%; Score 146; DB 6; Length 158;
Best Local Similarity 99.4%; Pred. No. 7.2e-37;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGTGTCACTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 59
DB 1 CGTGTATCTCACATGGCATCTACATGCTTGTTCACCGTTGGTTCATCGTCC 60

QY 60 CAAGCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 119
DB 61 CAAGCTTGCTATTCTTGAACCAAGAGATACCTACTCCCAACATCATCTTACTCAT 120

QY 120 GCAACTTCATGCAAAACGACATATGTTCTCGAAC 157

Db	121	GCACTTCATGCAACAGCACATATGTTTCTGAAC	158
RESULT 8	AC135206	137327 bp	DNA
LOCUS	Oryza sativa (japonica cultivar-group)	linear	PLN 16-APR-2003
DEFINITION	OJ1041F02, complete sequence.	chromosome 3 clone	
ACCESSION	AC135206		
VERSION	AC135206.3	GI:27596977	
KEYWORDS	HTG.		
SOURCE	Oryza sativa (japonica cultivar-group)		
ORGANISM	Oryza sativa (japonica cultivar-group)		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J. and Collura,K.		
TITLE	Rice Genomic Sequence		
JOURNAL	Unpublished		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (09-OCT-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Saski,C., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (06-NOV-2002) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Kim,H.-R., Rambo,T., Currie,J., Collura,K. and Thompson,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (11-JAN-2003) Arizona Genomics Institute, University of Arizona, 303 Forbes, Tucson, AZ 85721, USA		
REFERENCE	Wing,R.A., Yu,Y., Soderlund,C., Chen,M., Kim,H., Rambo,T., Saski,C., Henry,D., Thompson,S., Simmons,J., Thurmond,S.K. and Sun,S.		
TITLE	Direct Submision		
JOURNAL	Submitted (16-APR-2003) Clemson University Genomice Institute, Clemson University, 100 Jordan Hall, Clemson, SC 29634, USA		
COMMENT	On Jan 11, 2003 this sequence version replaced gi:24635891. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by more than one plasmid subclone; and the assembly was confirmed by restriction digest. There are transposons located at area 29427-32874 and at area 19945-33212. Region 70511-70536 is covered by a single subclone. Region 18678-18873 is a double stranded single subclone. Areas 69957-70761, 89836-90416, 98951-99009, 105316-105607, and 133052-133152 are covered by Monsanto only. The nucleotide sequence of this BAC clone was generated by combing SynGene, Monsanto and Arizona Genomics Institute sequencing data.		
FEATURES	Location/Qualifiers		
source	1. 137327		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:39947"		
	/chromosome="3"		
	/clone="OJ1041F02"		
	/note="(japonica cultivar-group)"		
gene	complement(638..9495)		
	/gene="OJ1041F02.1"		
	/note="unknown protein"		
CDS	complement(Join(638..838,1141..1276,3132..3244,3326..3421,3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,		
	5322..5411,5484..5699,5808..5942,6035..6538,6622..6837,		
	7548..7593,7679..7815,8196..8328,9221..9495))		
	/gene="OJ1041F02.1"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06844.1"		
	/db_xref="GI:29893590"		
	/translation="MAQSLLELLIOFLMPDNDARQAEQIRLARDPOVVALVHNLPTAKPRVQALAAVILRRKITSHPKLPRAKASIKQALISITTDHSHLRRAANVVSITAKVAPAGBPELPIFQCSQSFQEDHREVALILFSLSTIGITRQSHLDANOPTILKCLQDASRSRRIALKAAGSFIEVNDGDGVKIFRDFPSTILNVSQCLANGBEDVASTAFEPIDELIESPAPLGDVSRSIVQFELVCSNOELINIROAQIQTISLVKFEASFELKGLKIVIPILQVCCPLTATADGDGSDLAARSAEVALDTMAINPRHVPFPLERASVSPRINRYREAAVTGIVSGEGCHELKKKLECLRVLEALDOEOMRRAASPAFLGCPREHIQPRILSYSEVLICINALRDPDVEYKSYVALAACEDMGENTLPILDELMKLVNSLOGSPRNLETCMSAIGVAAABQAFMPAEVILVEMKGFMTVNDDELCAARATEVGIYAMAAGRAMETILPFIEALISGVLDSELEYTHGFPSNVAELIDDSFAQYLPHVVPVLAASCLDLDGSAVDIDADSVNGSEGSDDDVNDEPRVANSIWLTIILADILTAIRAIIPRAADVLEKOKDILDTMYNIYIKMRDDDKVVQAQCTSLADIVRDCGFALIEBYIRLADATILLRQSCCOQVSDGDDGDDIDHDEVMDAVSDLPAPFAKMGSYDPDPIFYKFDLSMKRAKSPHPQDITMYATIAEVAQMGAPISAVYDKMPLVILKELASSEATRNNAFCVEMCKNGGAALATKYGDILHGLHRLPADSEPDVADRDNAAGAIAMINVOQSIPLNDILPLVPDVINAAQVAVVSPNLSDEVKTVAAVASHLISVYEQMPILSALPPAHANALASFANR"		
gene	12372..15503		
	/gene="OJ1041F02.2"		
	/note="unknown protein"		
CDS	Join(12372..12757,12858..13025,13130..13295,13716..13826,13982..14130,14224..14354,14459..14558,15197..15503)		
	/gene="OJ1041F02.2"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06845.1"		
	/db_xref="GI:29893591"		
	/translation="MPQQSECCRHGVLGEBAAVGGSEEPFVPLRKRSVDGKASTAPRICIWCDEGAEADITCDIYAAPLRSCAKAMPAPLPFMMTPPPRQORDGEBAAPRGEAIRKGNHSDISLMLNQLGISYVSXKSLPPOKLAASDPDRKWTFFPEGSKETPPHNSVDPRWKDQCPAVFRHLKRLGVDPAEYVLAICGDTLRELAISPKGSCFRTTODDRMKTIVKSEYKVLIRMRSLYEHVROVKSLTLRFVTHICAKGCPKIHRRFDLKGSSHGRTIDTKRKIDETTLTKDLDQYAFRLQRFYBELMKQIONDCTPLETQGLIDVSLGLGVHFRNDYSVKIGISQHLAPKSTGRKSESGSSFCELCFVSEGCORDDIDRKEPTQLGIMNPQAEKSSKILDNPLNRRHLFTTPSGGSCCVLFPGLIDLDQYDITKLEHAYKSFQVNPDIYSAVDPKLRSRRFQIRFVFIKEQ"		
repeat_region	16734..16878		
	/note="putative MITE, GalTjn/Gaigln-like"		
repeat_region	19263..19454		
	/note="putative transposon, CACTA, En/Spm-like"		
gene	21819..24660		
	/gene="OJ1041F02.3"		
	/note="hypothetical protein"		
CDS	Join(21819..22007,22096..22206,23772..23996,24523..24660)		
	/gene="OJ1041F02.3"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="AAP06846.1"		
	/db_xref="GI:29893592"		
	/translation="MATASDGGSGGGEORRLISTPKEGERTIAPRRPDGTLRKAIRIRAGYPROEVALIYSGKAGQMKSGSDVPGYDPALDAPKTKAKAKNRKRRQKQAS TTNDGKGLHIEDGATDNPKDNDVSTKQISGLAISLSLVATSSYTDATDNRNSAS SAPDIDKKIRALKKKIRLAEAOVGDPENLKPEOLEKMKIKEGMKELKLENNKSSPAAS"		
gene	complement(26563..30575)		
	/gene="OJ1041F02.4"		
	/note="unknown protein"		
CDS	complement(Join(26563..26772,27082..27288,27430..27537,27881..27940,28051..28126,28327..28484,28616..28678,28790..28874,29560..30575))		
	/gene="OJ1041F02.4"		
	/codon_start=1		
	/product="unknown protein"		
	/protein_id="AAP06847.1"		
	/db_xref="GI:29893593"		
	/translation="WGPPAAPSPSSSSSGGSSRRRRRLRLERRNNAKHIGYDASNFCAYP		

RESULT	11
BX950854/c	
LOCUS	BX950854
DEFINITION	Dnao ret1o clone CH211-125M22, WORKING DRAFT SEQUENCE, 5 unordered places.
ACCESSION	BX950854
VERSION	BX950854.5 GI:52694542

KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 174806)
AUTHORS McLaren,S.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Sep 25, 2004 this sequence version replaced gi:42821090.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zc125M2
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 172622 bases at least Q40
Consensus quality: 173174 bases at least Q30
Consensus quality: 173717 bases at least Q20
Insert size: 174406; sum-of-contigs
Insert size: 182163; 2.3% error; agarose-fp
Quality coverage: 11.29x in Q20 bases; sum-of-contigs Quality coverage: 11.03x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 74620: contig of 74620 bp in length
* 74621 74720: gap of 100 bp
* 74721 133368: contig of 58648 bp in length
* 133369 133468: gap of 100 bp
* 133469 150793: contig of 17325 bp in length
* 150794 150893: gap of 100 bp
* 150894 172651: contig of 21758 bp in length
* 172652 172751: gap of 100 bp
* 172752 174806: contig of 2055 bp in length.
*
Location/Qualifiers
1. 174806
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-125M22"
/clone_lib="CHORI-211"
1. 74620
/note="assembly_fragment:02018
fragment_chain:1
clone_end:SP6
vector_side:left"
74721. 133368
/note="assembly_fragment:00796
fragment_chain:1"
133469. 150793
/note="assembly_fragment:00037
fragment_chain:1"
150894. 172651
/note="assembly_fragment:00366
fragment_chain:1"
172752. 174806
/note="assembly_fragment:00028
clone_end:T7
vector_side:right"

ORIGIN
Query Match 22.2%; Score 34.8; DB 2; Length 174806;
Best Local Similarity 52.8%; Pred. No. 2.7;
Matches 75; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 15 TGGCATCTACATGGTTGTCACACGCTGCTTGTTCACATGTCACAGCTTGCCTATT 74
DB 162638 TGGGTTGTGTCATCTTACACCAAGCTGGGGTGTTAAACACACCGCTTTAAAGT 162639
QY 75 CTGAACCAAGAGATACCTACTCCCAACATTCATCTTACTGATGCAACTTCATGCA 134
DB 162638 AAGTAACATCTACTTAAAGCCCTAAATATTTCATTTAAATCTGAATCTCAGTTTAA 162679
QY 135 ACACGCACATATGTTCTCGAA 156
DB 162578 ACTCCACATTTCTCATTTAAA 162557
RESULT 12
AC104134
LOCUS AC104134 110127 bp DNA linear PRI 20-MAR-2002
DEFINITION Homo sapiens BAC clone RP11-525L16 from 2, complete sequence.
ACCESSION AC104134
VERSION AC104134.4 GI:19551199
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 110127)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 110127)
AUTHORS Tomlinson,C., Haakenson,W. and Doeberer,A.
TITLE The sequence of Homo sapiens BAC clone RP11-525L16
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 110127)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 110127)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 110127)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Mar 20, 2002 this sequence version replaced gi:18677630.
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplena@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0525L16

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoeogawa, K., Moon, P.Y., Zhao, B., Frengsen, E., Tateo, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.reagen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-554H10, 2000 bp overlap; the clone sequenced to the right is RP11-450B3, 2000 bp overlap. Actual start of this clone is at base position 151996 of RP11-554H10.

FEATURES

source

1..110127
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-525L16"
/clone_1bp="RP11-11"
701..598
/rpt_family="Alu"
repeat_region
999..1297
/rpt_family="Alu"
repeat_region
1410..1486
/rpt_family="L1"
repeat_region
1564..1678
/rpt_family="Alu"
repeat_region
1684..1722
/rpt_family="AT_rich"
repeat_region
2338..2456
/rpt_family="MIR"
repeat_region
4039..4332
/rpt_family="Alu"
repeat_region
4412..4532
/rpt_family="Alu"
repeat_region
4669..4846
/rpt_family="MIR"
repeat_region
5878..6173
/rpt_family="Alu"
repeat_region
6385..6584
/rpt_family="MIR"
repeat_region
6929..7121
/rpt_family="MIR"
repeat_region
7167..7312
/rpt_family="MIR"
repeat_region
7953..8202
/rpt_family="Alu"
repeat_region
8537..8577
/rpt_family="L2"
repeat_region
8585..8756
/rpt_family="L1"
repeat_region
8768..8888
/rpt_family="Alu"
repeat_region
9005..9219
/rpt_family="L2"

repeat_region 9272..9838
/rpt_family="L1"
repeat_region 9849..9976
/rpt_family="Alu"
repeat_region 10124..10411
/rpt_family="Alu"
repeat_region 10483..10759
/rpt_family="L2"
repeat_region 11215..11264
/rpt_family="(TA)n"
repeat_region 11362..11476
/rpt_family="Alu"
repeat_region 11478..11776
/rpt_family="Alu"
repeat_region 12314..12607
/rpt_family="Alu"
repeat_region 12778..12943
/rpt_family="MalR"
repeat_region 12954..12979
/rpt_family="Alu"
repeat_region 14391..14655
/rpt_family="Alu"
repeat_region 14670..14939
/rpt_family="MalR"
repeat_region 15435..15449
/rpt_family="L2"
repeat_region 15450..15730
/rpt_family="Alu"
repeat_region 15731..15824
/rpt_family="L2"
repeat_region 15837..15915
/rpt_family="MalR"
repeat_region 16150..16214
/rpt_family="MalR"
repeat_region 16224..16382
/rpt_family="MER1_type"
repeat_region 16533..16747
/rpt_family="MIR"
repeat_region 16889..17116
/rpt_family="L2"
repeat_region 17271..17343
/rpt_family="L1"
repeat_region 17384..17534
/rpt_family="MIR"
repeat_region 17545..17843
/rpt_family="Alu"
repeat_region 17995..18056
/rpt_family="L2"
repeat_region 18057..18074
/rpt_family="GA-rich"
repeat_region 18075..18379
/rpt_family="Alu"
repeat_region 18380..18404
/rpt_family="GA-rich"
repeat_region 18889..19067
/rpt_family="MER1_type"
misc_feature 19296..19302
/note="match to EST BE843354 (NID:G10275732)"
repeat_region 19593..19709
/rpt_family="Alu"
repeat_region 19890..20198
/rpt_family="Alu"
repeat_region 20225..20698
/rpt_family="L1"
repeat_region 20737..20769
/rpt_family="AT_rich"
repeat_region 20887..21186
/rpt_family="Alu"
misc_feature 21350..21369
/note="match to EST BG571642 (NID:G13579295)"
repeat_region 21357..21652
/rpt_family="Alu"
repeat_region 21653..21957


```

repeat_region      22328..22636
                    /rpt_family="Alu"
Query Match      22.0%; Score 34.6; DB 9; Length 110127;
Match Local Similarity 54.3%; Pred. No. 3.1;
Base 70; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

OY 3 TGTCACTCAGATGCGATCTACATGCTGTTCACCGCTGCTGTTCATGTCATGCA 62
DB 8763 TGGGATTACAGAGGTAGCCACGTGACCTGCGCACTTCTTTCTCTCTCCCA 87692

OY 63 GCCTTGCTTATTGTAACCAAGAGATCTACTTCCCAACATTCATCTACTATGCA 122
DB 87693 TCTTCCCTTATCTTACCTCTGTTCATCTACTCAGAACTCACAATTCCTCATACC 87752

OY 123 ACTTCGATG 131
DB 87753 ACCTCCAG 87761

```

```

RESULT 13
AC136896      214946 bp      DNA      linear      PRI 21-MAR-2003
LOCUS      Homo sapiens chromosome 15, clone RP11-86001, complete sequence.
DEFINITION      AC136896
ACCESSION      AC136896
VERSION      AC136896.6
KEYWORDS      GI:29135656
SOURCE      HTG.
ORGANISM      Homo sapiens (human)

```

```

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      1 (bases 1 to 214946)
JOURNAL      Birren,B., Nusbaum,C. and Lander,E.
REFERENCE      Homo sapiens chromosome 15, clone RP11-86001
AUTHORS      Unpublished
              2 (bases 1 to 214946)
              Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
              Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
              Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
              Cook,A., Cooke,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
              Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
              Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
              Hago,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
              Matthews,C., McCarthy,M., Meldrim,J., Menus,L., Mihova,T.,
              Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
              Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
              Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
              Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R.,
              Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
              Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
              Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL      Submitted (09-NOV-2002) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS      3 (bases 1 to 214946)

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS      Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
              Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
              Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
              Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
              Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
              Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
              Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,
              Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
              Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
              Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
              Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
              O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
              Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
              Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,

```

```

TITLE
JOURNAL      Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
              Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
              Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
              Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL      Submitted (23-JAN-2003) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS      4 (bases 1 to 214946)
              Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
              Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
              Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
              Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
              Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
              Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
              Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hago,B.,
              Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
              Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
              Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
              Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
              Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
              Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
              O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
              Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
              Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
              Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
              Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
              Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
              Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL      Submitted (21-MAR-2003) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT      On Mar 21, 2003 this sequence version replaced gi:2877278.
              All repeats were identified using RepeatMasker:
              Smit, A.P.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: 127824
              Center clone name: 860_O_1
              -----
FEATURES
source      1..214946
              location/Qualifiers
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="15"
              /map="15"
              /clone="RP11-86001"
              /clone_lib="RPET-11 Human Male BAC"
              /complement(1..362)
              /rpt_family="HML1"
              /complement(823..1431)
              /rpt_family="L1MC5"
              913..922
              /note="<30 qual SNGL region"
              /complement(2941..3295)
              /rpt_family="MER31B"
              /complement(3307..3311)
              /note="<30 qual SNGL region"
              3438..3476
              /note="single clone coverage"
              /note="single clone coverage"
              3546..3586
              /rpt_family="GA-rich"
              3601..4006
              /rpt_family="MT1A1"
              complement(4008..4166)
              /rpt_family="MIR"
              complement(4580..4711)
              repeat_region

```



```

repeat_region      /rpt_family="L3"
4731..4751
/rpt_family="AT_rich"
repeat_region      complement(4758..4993)
/rpt_family="MIR"
repeat_region      complement(5336..5464)
/rpt_family="L3"
repeat_region      5935..6412
/rpt_family="MER39"
repeat_region      complement(6466..7069)
/rpt_family="MLT1E"
repeat_region      complement(8633..8713)
/rpt_family="MIR"
repeat_region      8768..9244
/rpt_family="L1M2"
repeat_region      complement(9299..9498)
/rpt_family="MIR"
repeat_region      10252..10587
/rpt_family="MLT1A0"
repeat_region      12232..12619
/rpt_family="L1M5"
repeat_region      complement(13183..13354)
/rpt_family="MIR"
repeat_region      14074..14500
/rpt_family="L1M3"
repeat_region      15070..15201
/rpt_family="FLAM_C"
repeat_region      complement(16057..18621)
/rpt_family="L1M3"
repeat_region      complement(18678..19377)
/rpt_family="L1M1"
repeat_region      complement(19375..20142)
/rpt_family="L1M3"
repeat_region      22615..22919
/rpt_family="A1USg"
repeat_region      complement(25101..25206)
/rpt_family="HAL1"
repeat_region      complement(25691..25925)
/rpt_family="HAL1"
repeat_region      25996..26033
/rpt_family="CAGA)n"
repeat_region      complement(26059..26582)
/rpt_family="MLT2B3"
repeat_region      26645..27455
/rpt_family="MER21B"
repeat_region      complement(27582..27833)
/rpt_family="HAL1"
repeat_region      complement(28185..28759)
/rpt_family="LTR68"
repeat_region      complement(29229..29603)
/rpt_family="THE1A"
repeat_region      complement(29771..30036)
/rpt_family="A1uB"
repeat_region      30056..30083
/rpt_family="TPTA)n"
repeat_region      complement(30090..30369)
/rpt_family="A1uB"
repeat_region      31042..32585
/rpt_family="L1M4"
repeat_region      32586..32654
/rpt_family="TCTA)n"
repeat_region      32655..33163
/rpt_family="L1M4"
repeat_region      33164..33473
/rpt_family="A1uY"
repeat_region      33474..33683
/rpt_family="L1M4"
repeat_region      33686..33751
/rpt_family="MLT1D"
repeat_region      36941..37015
/rpt_family="MIR"
repeat_region      complement(38001..38449)
/rpt_family="L1MD1"

```

```

repeat_region      38453..38855
/rpt_family="MLT1D"
repeat_region      complement(38856..39656)

Query Match      21.9%; Score 34.4; DB 9; Length 214946;
Best Local Similarity 54.8%; Pred. No. 3.7;
Matches 68; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy      5 TCATCTCATGCGCATCTACATGCTGTTCACCGTTCGTCTGTTCATCGTCCAGC 64
|||||
Db      35745 TCATCTCATGCGCGAGAAAAATCATTTGACAAATTAGGCATTTCTTCATGATGTAAAC 35804
|||||

Qy      65 CTGGCCATTTCTGACCAAGAGATACCTACTGCCAAACATTCATCTACTATGCAAC 124
|||||
Db      35805 ATCCAACTATGTAAGAAAGAAAGTGTCTTTTAAACCAATTAATGTAACCTATGTAAGA 35864
|||||

Qy      125 TTCC 128
Db      35865 ATCC 35868

```

```

RESULT 14
CR847532/c      141810 bp      DNA      linear      HTG 03-OCT-2004
LOCUS
DEFINITION
Danto rerio clone DKEXP-86C9, *** SEQUENCING IN PROGRESS ***, 8
unordered pieces.
ACCESSION
CR847532
VERSION
CR847532.2 GI:53755907
HTG; HTGS PHASE1.
KEYWORDS
Danto rerio (zebrafish)
SOURCE
Danto rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danto.
1 (bases 1 to 141810)
McLay, K.

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

```

COMMENT

```

Direct Submission
Submitted (02-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
fish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 3, 2004 this sequence version replaced gi:53145936.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zK086C9
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 13909 bases at least Q40
Consensus quality: 140415 bases at least Q30
Consensus quality: 140735 bases at least Q20
Insert size: 14110; sum-of-contigs
Insert size: 140600; 2.5% error; agarose-fp
Quality coverage: 8.52x in Q20 bases; sum-of-contigs Quality
coverage: 8.64x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 19172: contig of 19172 bp in length
* 19173 19272: gap of 100 bp
* 19273 46769: contig of 27497 bp in length
* 46770 46869: gap of 100 bp
* 46870 52415: contig of 5546 bp in length
* 52416 52515: gap of 100 bp

```


*	52825	64328:	cont'g of 11813 bp in length
*	64429	64428:	gap of 100 bp
*	64429	96228:	cont'g of 31796 bp in length
*	96425	96382:	gap of 100 bp
*	96425	111956:	cont'g of 25632 bp in length
*	12157	122056:	gap of 100 bp
*	122057	131303:	cont'g of 9276 bp in length
*	131333	131433:	gap of 100 bp
*	131333	118180:	cont'g of 10378 bp in length

```

UNES
source
Location/Qualifiers
1. .141810
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/cclone="DKRYP-86C9"
/cclone_lib="Daniokeyplot"
1. .13172
/note="assembly_fragment:00300
fragment_chain:1"
19273. .46769
/note="assembly_fragment:00469
fragment_chain:1"
46870. .52515
/note="assembly_fragment:00017
fragment_chain:1"
52516. .64328
/note="assembly_fragment:00167
fragment_chain:1"
64429. .95624
/note="assembly_fragment:00976
fragment_chain:2"
96325. .121956
/note="assembly_fragment:01302
fragment_chain:2"
122057. .131332
/note="assembly_fragment:00056.0"
131433. .141810
/note="assembly_fragment:00703"

```

Query Match	Similarity	21.5%	Score 33.9	DB 2	Length 141810
Best Local Alignment	52.5%	Pred. No. 5.8			
Matches 74	Conservative	0	Mismatches	67	Indels
					Gaps
					0
Oy	12	ACATGGCATACTACATGCTTGTTCACCGCTTCGCTTGTTGCATTCGCGCAAGCCTTGCC	71		
Db	41846	ACAGTGAACCCCTACATCTTGTGCTCAATGSCATATTGTGCAAGTACAGAGGTTGGG	41787		
Oy	72	ATTCTGAACCAAGAGATACCTTATCCCAACAATCATCTTACTCATGCACTTCAC	131		
Db	41786	AAGTGGGAAATGGGGATCATCTGCTCCTGAAAAATGAAAAATCAACCAAGAAACT	41722		
Oy	132	CAACACGCACATATGTTCC	152		
Db	41726	TCAAAAAAACTTATTAGTCC	41706		
RESULT 15					
LOCUS	BX640577	214178 bp	DNA	linear	VRT 10-JUL-2004
DEFINITION	zebrafish DNA sequence from clone DKEX-117014 in linkage group 14,				
ACCESSION	BX640577				
VERSION	BX640577.17	GI:50080105			
KEYWORDS	HTG.				
SOURCE	Danio rerio				
ORGANISM	Danio rerio (zebrafish)				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Actinopterygii; Neopterygii; Teleostei; Ostariophysi;				
TITLE	Cypriniformes; Cyprinidae; Danio.				
	1 (bases 1 to 214178)				
	McGuire, S.				
	Direct Submission				

JOURNAL Submitted (10-JUL-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT On Jul 8, 2004 this sequence version replaced g1:43653408.

Submitted (10-JUN-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk; Clone requests: clonerequest@sanger.ac.uk
On Jul 8, 2004 this sequence version replaced g1:4959408.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhong Rao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_reio/fishmark.shtml DKEX-117014 is from a Zebrafish BAC library

RECTOR: pIndiOBAC-5.

```

FEATURES
  source
    Location/Qualifiers
      1..2141178
        /organism="Danio rerio"
        /mol_type="genomic DNA"
        /db_xref="taxon:7955"
        /clone="DKEX-117014"
        /clone_1fb="DanioKey"

```

	Query Match	21.5%	Score 33.8	DB 5	Length 214178
	Best Local Similarity	52.5%	Pred. No. 5.9		
	Matches 74	Conservative 0	Mismatches 67	Indels 0	Gaps 0
Qy	12	ACATGGCATCTACTCATGCTTTGTTCAACGCTGGTCTTTGTTTCATCGTCCAAAGCTTGCGCT	71		
Db	58301	ACAGTGAACCCCTTACATCCTTTGGTCCCTCAATGCAATATTGTACAAAGTGTACAGAGTTGGCT	58360		
Qy	72	ATTCTGAACCAAGAGGATACCTACTCCAAACATCATCTTACTCATGCAATTTCCAGT	131		
Db	58361	AAGTGGGAATATGGGATCATTTGCTCTCTAATAAATGATAAATTAATTCACACAAAGAACTT	58420		
Qy	132	CAAAACGACATATGTTTCC	152		
Db	58421	TCAAAAAAACTATATTAGTCC	58441		

Search completed: September 16, 2005, 03:01:19
Job time : 531.823 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:10:17 ; Search time 76.0101 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311

Perfect score: 157
Sequence: 1 cgcgcacatccacacatgcat.....cgcacataatgcttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Genesegq16Dec04:*
2: genesegq19808:*
3: genesegq19908:*
4: genesegq20008:*
5: genesegq2001a:*
6: genesegq2001b:*
7: genesegq2002a:*
8: genesegq2002b:*
9: genesegq2003a:*
10: genesegq2003b:*
11: genesegq2003c:*
12: genesegq2004a:*
13: genesegq2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	255	AAH76340	Aah76340 Z. mays M
2	157	100.0	1394	AAH76340	Aax07408 Z. mays M
3	157	100.0	1394	AAH76340	Aax07409 Z. mays M
4	157	100.0	1394	AAH76332	Aah76332 Z. mays M
5	157	100.0	1394	AAH76332	Aah76333 Z. mays M
6	146	93.0	158	AAH76334	Aah76334 Z. mays M
7	50	31.8	50	AAH76336	Aah76336 Z. mays M
8	40	25.5	40	AAH76337	Aah76337 Z. mays M
9	32.4	20.6	3267	ADK52131	Adk52131 Mouse atc
10	31.6	20.1	2352	ACA48402	ACA48402 Prokaryot
11	31.4	20.0	15734	ABN80329	Abn80329 Human che
12	30.6	19.5	64415	AAF22279	Aaf22279 BAC conta
13	30.6	19.5	79467	ADA02717	Ada02717 Mouse nfa
14	30.6	19.5	79467	ADA02717	Ada02717 Mouse nfa
15	30.6	19.5	79467	ADA02717	Ada02717 Mouse nfa
16	30.4	19.4	52001	ADP85965	Adp85965 Mouse nfa
17	30.2	19.2	669	ADA07076	Ada07076 Human mit
18	30.2	19.2	1266	ADA07076	Ada07076 Human mit
19	30.2	19.2	110000	ADA07076	Ada07076 Human mit
20	30	19.1	30	AAH76335	Aah76335 Z. mays M

C	21	30	19.1	378	10	ADD33403	Add33403 Mouse mit
C	22	30	19.1	378	10	ADD33404	Add33404 Mouse mit
C	23	29.8	19.0	373	4	AAH13328	Aah13328 Human cdn
C	24	29.6	18.9	1185	10	ADP00796	Adp00796 Bacterial
C	25	29.4	18.7	110000	3	AAF22303_2	AAF22303_2 of
C	26	29.2	18.6	1475	9	AAI62902	AAI62902 Vernonia
C	27	29.2	18.6	110000	2	AAI72063_06	AAI72063_06 of
C	28	28.8	18.3	1500	8	ADA68284	Ada68284 Arabidops
C	29	28.6	18.2	756	13	AD557339	Ad557339 Bacterial
C	30	28.6	18.2	778	4	AA532892	Aa532892 Human gen
C	31	28.6	18.2	778	4	AA532893	Aa532893 Human gen
C	32	28.6	18.2	1043	10	ADD71973	Add71973 Human xrt
C	33	28.6	18.2	1860	6	ABN7847	Abn7847 Streptoco
C	34	28.6	18.2	51615	11	ACN45162	Act45162 Human gen
C	35	28.6	18.2	110000	13	ABD32806_5	ABD32806_5 of
C	36	28.6	18.2	164702	8	ACF62730	Acf62730 Cancer ba
C	37	28.6	18.2	164702	8	ADB80845	Adb80845 MRP1 base
C	38	28.6	18.2	164702	10	ADB87934	Adb87934 Human UGT
C	39	28.6	18.2	164702	10	ADB96917	Adb96917 Human MDR
C	40	28.6	18.2	164702	10	ADB92108	Adb92108 Human MDR
C	41	28.4	18.1	1478	3	AAF15843	Aaf15843 Human pro
C	42	28.4	18.1	1715	8	ABA00762	Abao0762 PCR rela
C	43	28.4	18.1	1791	12	ADQ86652	Adq86652 Human tum
C	44	28.4	18.1	1791	12	ADQ85536	Adq85536 Human tum
C	45	28.4	18.1	1791	12	ADQ84654	Adq84654 Human tum

ALIGNMENTS

RESULT 1
ID AAH76340 standard; DNA; 255 BP.
XX
AC AAH76340;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 promoter fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; promoter; de.
XX
OS Zea mays.
XX
PN WQ200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Example 5; Fig 8; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a Z. mays Ms45 promoter fragment


```
XX
SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.3e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 60
DB 15 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 74
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 120
DB 75 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 134
QY 121 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB 135 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 171

RESULT 2
AAK07408 standard; DNA; 1394 BP.
AC AAK07408;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
KM Ms45; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; maize; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN WO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR MPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 2; Page 22-23; 39pp; English.
XX
CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 60
DB 1155 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 1214
```

```
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 120
DB 1215 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 1311

RESULT 3
AAK07409 standard; DNA; 1394 BP.
AC AAK07409;
XX
DT 08-JUN-1999 (first entry)
XX
DE Zea mays Ms45 male tissue-preferred regulatory region.
XX
KM Ms45; male; tissue-preferred; regulatory region; plant cells;
KW plant tissue; differentiated; hybrid seed; fertility; ss.
XX
OS Zea mays.
XX
PN WO9859061-A1.
XX
PD 30-DEC-1998.
XX
PF 19-JUN-1998; 98WO-US012895.
XX
PR 23-JUN-1997; 97US-00880499.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
XX
DR MPI; 1999-105628/09.
XX
PT New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
XX - useful in mediating plant fertility, especially hybrid seed production.
XX
PS Claim 3; Page 23-24; 39pp; English.
XX
CC The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolB and diptheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 2; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 60
DB 1155 CGTGTATCTCATGAGCATGATGCTGTTGCAACCGTTGCTGTTGCATGCTGC 1214
QY 61 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 120
DB 1215 AAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCATCTTACTCAG 1274
QY 121 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 157
DB 1275 CAACCTTCATGCAACACGACATATGTTCTCTGAAC 1311
```



```
RESULT 4
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
XX
AC AAH76332;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 46; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAGATGCGATGACTGCTTGTTCACCGCTTGTTCATGCTGC 60
DB 1155 CGTGCATCTCAGATGCGATGACTGCTTGTTCACCGCTTGTTCATGCTGC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311
RESULT 5
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
XX
AC AAH76333;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
```

```
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
DR WPI; 2001-514772/56.
XX
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX
PS Claim 4; Page 47; 50pp; English.
XX
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 157; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGTGCATCTCAGATGCGATGACTGCTTGTTCACCGCTTGTTCATGCTGC 60
DB 1155 CGTGCATCTCAGATGCGATGACTGCTTGTTCACCGCTTGTTCATGCTGC 1214
QY 61 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG 120
DB 1215 AAGCCTTGCTATTCTGAACCAAGAGATACCTACTCCCAACAATCCATCTTACTCATG 1274
QY 121 CAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 157
DB 1275 CAACTTCATGCAAAACGCGACATATGTTTCTGTAAC 1311
RESULT 6
AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
XX
AC AAH76334;
XX
XX
DT 29-OCT-2001 (first entry)
XX
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX
XX
KW hybrid seed; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
```



```

XX 13-FEB-2001; 2001WO-US004527.
PF
XX 15-FEB-2000; 2000US-00504487.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI
XX WPI; 2001-514772/56.
DR
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
SQ
XX
XX Query Match          93.0%; Score 146; DB 5; Length 158;
XX Best Local Similarity 99.4%; Pred. No. 5.4e-39;
XX Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX QY 1 CGTGTACATCTACATGCGATCTACATGCTGTTCACCGTTCGTC-TGTTCATCGTC 59
XX DB 1 CGTGTACATCTACATGCGATCTACATGCTGTTCACCGTTCGTCATCGTC 60
XX QY 60 CAAGCCTTGCTTCTGAAACCAAGAGATACCTACCTCCCAACATCTTACTCAT 119
XX DB 61 CAAGCCTTGCTTCTGAAACCAAGAGATACCTACCTCCCAACATCTTACTCAT 120
XX QY 120 GCAACTTCCATGCAACACGACATATGTTCTCTGAAAC 157
XX DB 121 GCAACTTCCATGCAACACGACATATGTTCTCTGAAAC 158
XX
XX RESULT 7
XX AAH76336
XX ID AAH76336 standard; DNA; 50 BP.
XX AC AAH76336;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX PF
XX 15-FEB-2000; 2000US-00504487.
XX PR
XX (PION-) PIONEER HI-BRED INT INC.
XX PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX PI

```

```

XX WPI; 2001-514772/56.
DR
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
SQ
XX
XX Query Match          31.8%; Score 50; DB 5; Length 50;
XX Best Local Similarity 100.0%; Pred. No. 7.2e-07;
XX Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 75 CTGAACCAAGAGATACCTACTCTCCCAACATCTTACTCATGCAAC 124
XX DB 1 CTGAACCAAGAGATACCTACTCTCCCAACATCTTACTCATGCAAC 50
XX
XX RESULT 8
XX AAH76337
XX ID AAH76337 standard; DNA; 40 BP.
XX AC AAH76337;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX PN
XX 23-AUG-2001.
XX PD
XX 13-FEB-2001; 2001WO-US004527.
XX PF
XX 15-FEB-2000; 2000US-00504487.
XX PR
XX (PION-) PIONEER HI-BRED INT INC.
XX PA
XX Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX PI
XX WPI; 2001-514772/56.
XX DR
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 14; Page 32; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The

```


CC present sequence represents a DNA fragment upstream of the TATA box of a
CC Z. mays Me45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
Query Match 25.5%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 124
DB 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
RESULT 9
ADK52131
ID ADK52131 standard; cDNA; 3267 BP.
XX
AC ADK52131;
XX
DT 20-MAY-2004 (first entry)
XX
DE Mouse atopic-dermatitis/psoriasis-associated EST #4.
XX
KM Mouse; ss; EST; atopic dermatitis; psoriasis; dermatological;
XX anti-inflammatory; antipsoriatic; rash; expressed sequence tag.
XX
OS Mus musculus.
XX
PN W02004016785-A1.
XX
PD 26-FEB-2004.
XX
PF 06-AUG-2003; 2003MO-JP009999.
XX
PR 06-AUG-2002; 2002JP-00229319.
XX
PR 14-MAY-2003; 2003JP-00136544.
XX
PA (GENO-) GENOX RES INC.
PA (UOJU-) UNIV JUNTENDO.
XX
PI Itoh M, Ogawa K, Shinagawa A, Sudo H, Ogawa H, Ra C;
PI Mitsuihara K;
XX
DR WPI; 2004-214514/20.
XX
PT Detecting atopic dermatitis or psoriasis comprises assaying levels of
PT expression of an indicator gene at a rash site and non-rash site of a
PT person with atopic dermatitis or psoriasis.
XX
PS Claim 20; SEQ ID NO 164; 484bp; Japanese.
XX
XX The invention relates to detecting atopic dermatitis or psoriasis
CC comprising assaying the levels of expression of an indicator gene at a
CC rash site and non-rash site of a person with atopic dermatitis or
CC psoriasis, comparing these levels with those of a healthy person, and
CC determining that if the levels of indicators are higher or lower, then
CC this indicates the disease. Also included are a reagent for detecting
CC atopic dermatitis or psoriasis, a kit for screening for treatments, a
CC transgenic non human vertebrate animal models for the diseases, an agent
CC for inducing the diseases in mice and a DNA chip for assaying for the
CC indicator gene. The method is used for treatment, detection and animal
CC models for research of atopic dermatitis and psoriasis. The present
CC sequence is a Mouse atopic-dermatitis/ psoriasis-associated EST
CC (expressed sequence tag).
XX
SQ Sequence 3267 BP; 1029 A; 658 C; 590 G; 990 T; 0 U; 0 Other;
Query Match 20.6%; Score 32.4; DB 12; Length 3267;
Best Local Similarity 54.1%; Pred. No. 2.2;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
QY 27 TGGTGTTCACACGCTTGTGTCATGTCACAGCCTTGCTATTTCGACCAAGAG 86

DB 2774 TTTTTCCTACCTCCCTCTGATTTCTTTGACCTCATTCTTATTTGGCCTTTAAGTT 2833
QY 87 GATACCTACTCCCAACATCCATCTTACTCATGCAACTTCACGCAACACGCACATAT 146
DB 2834 GTTCTAGATTCCCAAGGAAGCCATTTTGACCTACAGTCAAGCTACTTAACACCTGAAAT 2893
QY 147 GT 148
DB 2894 GT 2895
RESULT 10
ACA48402/C
ID ACA48402 standard; DNA; 2352 BP.
XX
AC ACA48402;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #30059.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
OS Streptococcus mutans.
XX
PN W0200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002MO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
XX
PR 25-OCT-2001; 2001US-0342923P.
XX
PR 08-FEB-2002; 2002US-00072851.
XX
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
XX
DR P-PSDB; ABU44532.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation or
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 36272; 1766bp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibody; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of

CC cytokines, antibodies, and growth factors
XX
SQ Sequence 64415 BP; 18698 A; 13554 C; 13083 G; 18980 T; 0 U; 100 Other;
Query Match 19.5%; Score 30.6; DB 3; Length 64415;
Best Local Similarity 58.1%; Pred. No. 23;
Matches 57; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 65 CTGGCTATTCTGAAACCAAGAGATACCTACTCCCAACATCCATCTTACTATGCAAC 124
DB 7117 CTGGCTATTCTGAAATATATATATATTTGCTCTTAACACTAAACCTTAACCTACACC 7176
QY 125 TTCCATGCAACACGCGCATATGTTCTGAC 157
DB 7177 TTAAATCCCAACCTTAATCTAATTCCTTAAC 7209
RESULT 13
ADA02717
ID ADA02717 standard; DNA; 79467 BP.
XX
XX ADA02717;
XX
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Nfatc1 carcinoma associated gene, SEQ ID NO:1235.
XX
XX Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX KM gene; de.
XX
XX Mus sp.
XX OS
XX PN MO2003057146-A2.
XX PD 17-JUL-2003.
XX PF 26-DEC-2002; 2002MO-US041414.
XX PR 26-DEC-2001; 2001US-00035832.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX DR MPI; 2003-587068/55.
XX
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX
XX
PS Claim 1; SEQ ID NO 1235; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed murine CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;
Query Match 19.5%; Score 30.6; DB 9; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGCTCCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCCAT 110
DB 53779 TCCATGATGATCTTGCTTCTGCTGGAACAGCTGATCTTAATGACACACACACA 53838
QY 111 CTTATCTGATGCAACTTCCATGCAACACGCAATATGTTTC 151
DB 53839 CACACACACACACACACACACACACACAGCTTGCTGCTC 53879
RESULT 14
ADB72455
ID ADB72455 standard; DNA; 79467 BP.
XX
XX ADB72455;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Nfatc1 gene.
XX
XX mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
XX cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX KM
XX
XX Mus sp.
XX OS
XX PN MO2003008583-A2.
XX PD 30-JAN-2003.
XX PF 26-DEC-2001; 2001MO-US051291.
XX PR 02-MAR-2001; 2001US-00798586.
XX PR 23-OCT-2001; 2001US-00004113.
XX PR 08-NOV-2001; 2001US-00052482.
XX PR 30-NOV-2001; 2001US-00997722.
XX PR 20-DEC-2001; 2001US-00034650.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW, Engelhard EK;
XX DR MPI; 2003-239337/23.
XX
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
XX cancers, neoplasm, adenocarcinoma, and sarcomas.
XX
XX
PS Claim 1; SEQ ID NO 283; 2304pp; English.
XX
XX The invention relates to a novel recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the 660 sequences fully defined
XX in the specification. A polynucleotide of the invention has cytostatic
XX activity, and may have a use in gene therapy, or in a vaccine. The
XX recombinant nucleic acids and polypeptides are useful for treating
XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
XX sarcomas. The present sequence represents a mouse gene of the invention.
XX
XX
SQ Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;
Query Match 19.5%; Score 30.6; DB 10; Length 79467;
Best Local Similarity 56.4%; Pred. No. 25;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 51 TCCATGCTCCAGCCTTGCTTATTTGAAACCAAGAGATACCTACTCCCAACATCCAT 110
DB 53779 TCCATGATGATCTTGCTTCTGCTGGAACAGCTGATCTTAATGACACACACACA 53838

QY 111 CTTACTGATGCACTTCATGCAACGACGACATATGTTTC 151
 Db 53839 CACACACACACACACACACACACACACAGCTTGCGCTC 53879

RESULT 15

ADE95965
 ID ADE95965 standard; DNA; 79467 BP.

XX ADE95965;

XX 12-FEB-2004 (first entry)

XX Mouse Nfatc1 gene genomic DNA sequence.

XX cancer diagnosis; cancer treatment; carcinoma; cytostatic; gene therapy;
 KM lymphoma; breast cancer; prostate cancer; leukemia; ds; mouse; murine;
 KM Nfatc1.

XX Mus sp.

XX WO2003039484-A2.

XX 15-MAY-2003.

XX 08-NOV-2002; 2002WO-US036071.

XX 08-NOV-2001; 2001US-00052482.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;

XX MPI; 2003-441462/41.

XX New carcinoma associated nucleic acids and proteins, useful for screening
 PT drug candidates, or for diagnosing and treating carcinomas, e.g.
 PT lymphoma, breast cancer, prostate cancer or leukemia.

XX Claim 1; SEQ ID NO 223; 793pp; English.

CC This invention relates to novel recombinant nucleic acids for use in
 CC diagnosis and treatment of cancer, especially carcinomas, as well as the
 CC use of compositions in screening methods. The compositions of the
 CC invention may have cytostatic activity whilst the disclosed sequences may
 CC be useful for gene therapy. The carcinoma associated nucleic acids and
 CC proteins are useful for diagnosing and treating carcinomas, for example
 CC lymphoma, breast cancer, prostate cancer or leukemia, or for screening
 CC drug candidates or bioactive agents capable of binding to, or modulating
 CC the activity of, a carcinoma associated protein. The present sequence is
 CC the genomic DNA sequence of the mouse Nfatc1 gene which is a carcinoma
 CC associated gene of the invention.

XX Sequence 79467 BP; 18256 A; 19684 C; 20062 G; 20204 T; 0 U; 1261 Other;

Query Match 19.5%; Score 30.6; DB 10; Length 79467;

Best Local Similarity 56.4%; Pred. No. 25;

Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCCAGCCTTGCTTCTATTTGACCAAGAGATACCTTACTCCCAACATTCAT 110
 Db 53779 TCCACTAACTGATCTTGCTTCTCTGGAACAAGCTGATCTTAGTCACACACACACA 53838

QY 111 CTTACTGATGCACTTCATGCAACGACGACATATGTTTC 151

Db 53839 CACACACACACACACACACACACACACAGCTTGCGCTC 53879

Search completed: September 15, 2005, 21:33:10
 Job time : 81.0101 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 ; Search time 23.7481 Seconds
(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1155_1311

Perfect score: 157
Sequence: 1 cgcgcacatccacatgcat.....cgcacatgcttcctgaac 157

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/6C COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	157	100.0	1394	3 US-08-880-499-1	Sequence 1, Appl
2	157	100.0	1394	3 US-08-880-499-2	Sequence 2, Appl
3	30.4	19.4	385136	4 US-09-949-016-16073	Sequence 16073, A
4	29.6	18.9	1185	4 US-09-543-681A-1081	Sequence 1081, A
5	29.2	18.6	1830121	4 US-09-557-884-1	Sequence 1, Appl
6	28.8	18.3	64190	4 US-09-643-990A-1	Sequence 14712, A
7	28.8	18.3	64190	4 US-09-949-016-14712	Sequence 14713, A
8	28.8	18.3	64190	4 US-09-949-016-15502	Sequence 15502, A
9	28.8	18.3	64190	4 US-09-949-016-17447	Sequence 17447, A
10	28.8	18.3	64190	4 US-09-949-016-191201	Sequence 191201, A
11	27.8	17.7	601	4 US-09-949-016-17447	Sequence 15502, A
12	27.8	17.7	11490	4 US-09-949-016-14149	Sequence 14149, A
13	27.8	17.7	14952	4 US-09-949-016-16740	Sequence 16740, A
14	27.4	17.5	1413	4 US-09-710-279-1813	Sequence 1813, Ap
15	27.4	17.5	1413	4 US-09-134-001C-1413	Sequence 1413, Ap
16	27.4	17.5	1794	4 US-09-270-767-10489	Sequence 10489, A
17	27.4	17.5	2472	1 US-08-425-299A-2	Sequence 2, Appl
18	27.4	17.5	2926	4 US-09-710-279-3747	Sequence 3747, Ap
19	27.4	17.5	3014	4 US-09-710-279-3680	Sequence 3680, Ap
20	27.4	17.5	3599	4 US-09-710-279-4003	Sequence 4003, Ap
21	27.2	17.3	3445	4 US-09-949-016-2556	Sequence 2556, Ap
22	27.2	17.3	3588	4 US-09-566-821-23	Sequence 23, Appl
23	27.2	17.3	6751	1 US-07-882-925A-5	Sequence 5, Appl
24	27.2	17.3	6751	1 US-08-184-012C-5	Sequence 12703, A
25	27.2	17.3	8939	4 US-09-949-016-12703	Sequence 12703, A
26	27.2	17.3	9076	4 US-09-949-016-17368	Sequence 17368, A
27	27.2	17.3	48994	4 US-09-949-016-14091	Sequence 14091, A

28	27.2	17.3	60593	4 US-09-949-016-13779	Sequence 13779, A
29	27.2	17.3	62776	4 US-09-949-016-17576	Sequence 17576, A
30	27.2	17.3	100848	4 US-09-596-002-39	Sequence 39, Appl
31	27.2	17.3	158735	4 US-09-949-016-11989	Sequence 11989, A
32	27.2	17.3	158735	4 US-09-949-016-17130	Sequence 17130, A
33	27.2	17.3	879	3 US-09-134-001C-678	Sequence 678, App
34	27.2	17.2	44180	4 US-09-949-016-12088	Sequence 12088, A
35	27.2	17.2	47030	4 US-09-949-016-13037	Sequence 13037, A
36	27.2	17.2	47030	4 US-09-949-016-13038	Sequence 13038, A
37	27.2	17.2	47030	4 US-09-949-016-15039	Sequence 15039, A
38	27.2	17.2	47030	4 US-09-949-016-15040	Sequence 15040, A
39	27.2	17.2	194537	4 US-09-949-016-12928	Sequence 12928, A
40	27.2	17.2	201529	4 US-09-949-016-12740	Sequence 12740, A
41	26.8	17.1	265	4 US-09-513-999C-8379	Sequence 8379, Ap
42	26.8	17.1	271	3 US-08-946-026-9	Sequence 9, Appl
43	26.8	17.1	3114	3 US-08-946-026-12	Sequence 12, Appl
44	26.8	17.1	156942	4 US-09-949-016-12227	Sequence 12227, A
45	26.8	17.1	156950	4 US-09-949-016-15946	Sequence 15946, A

ALIGNMENTS

RESULT 1
US-08-880-499-1
Sequence 1, Application US/08880499
Patent No. 6037523
GENERAL INFORMATION:
APPLICANT: Alberson, Marc C.
APPLICANT: Fox, Tim W.
APPLICANT: Carl, Garmaat W.
APPLICANT: Hultman, Gary A.
APPLICANT: Kendall, Timmy L.
TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION
TITLE OF INVENTION: AND METHOD OF USING SAME
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,499
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Sweeney, Patricia A.
REGISTRATION NUMBER: 32,733
REFERENCE/DOCKET NUMBER: 0578
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-880-499-1
Query Match 100.0%; Score 157; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 5.6e-45;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 CTCGCAACATCATCTACTCA 118
DB 880 ATCCCTAGATGCTTTTATATA 857

RESULT 5

US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 12 ACATGCGACTACTGCTTGTTCACCGTGTCTTGTTCATCGTCAAGCTTGCCCT 71
DB 677275 ACTTTAAATTCAGCGCTGCTGTCTTATTCATTGCAATTAAGCAAGATAGCCCT 677216
QY 72 ATTCTGAACCAAGAGATCTACTCCCAACATCATCTACTCATGCACTTCATG 131
DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156
QY 132 CAACACGCA 141
DB 677155 CAACCAACCA 677146

RESULT 6
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann

Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS v6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 18.6%; Score 29.2; DB 4; Length 1830121;
Best Local Similarity 51.5%; Pred. No. 66;
Matches 67; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 12 ACATGCGACTACTGCTTGTTCACCGTGTCTTGTTCATCGTCAAGCTTGCCCT 71
DB 677275 ACTTTAAATTCAGCGCTGCTGTCTTATTCATTGCAATTAAGCAAGATAGCCCT 677216
QY 72 ATTCTGAACCAAGAGATCTACTCCCAACATCATCTACTCATGCACTTCATG 131
DB 677215 ATGCTGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 677156
QY 132 CAACACGCA 141
DB 677155 CAACCAACCA 677146

RESULT 7
US-09-949-016-14712/c
; Sequence 14712, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 14712
;; LENGTH: 64190
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-14712

Query Match 18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 55 TCGTCCAGCCTTCTATTTGACCAAGAGATACCTACCCAAACATCCATCTTA 114
DB 355 TCATCCAAATAGTTCCACCTACTATCATCATGATGGAACGTAGTACTATCATCAAGCCAAATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAACT 284

RESULT 8
US-09-949-016-14713/c
;; Sequence 14713, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 14713
;; LENGTH: 64190
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-14713

Query Match 18.3%; Score 28.8; DB 4; Length 64190;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 55 TCGTCCAGCCTTCTATTTGACCAAGAGATACCTACCCAAACATCCATCTTA 114
DB 355 TCATCCAAATAGTTCCACCTACTATCATCATGATGGAACGTAGTACTATCATCAAGCCAAATATT 296
QY 115 CTCATGCAACTT 126
DB 295 CTCATGTAACT 284

RESULT 9
US-09-949-016-15502
;; Sequence 15502, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 15502
;; LENGTH: 11022
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-15502

Query Match 17.8%; Score 28; DB 4; Length 11022;
Best Local Similarity 63.2%; Pred. No. 22;
Matches 43; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 15 TGGCATCTACTACTGCTTGTTCAACCGTCTCTTGTCATGTCACAGCTTGCTATT 74
DB 4432 TGGCTTATTTCATGGGTGTTTATGGGTTAATGTTGATCATCATTTAGCTTGTCACAG 4491
QY 75 CTGAACCA 82
DB 4492 CTGAATCA 4499

RESULT 10
US-09-949-016-17447
;; Sequence 17447, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 17447
;; LENGTH: 41815
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-17447

Query Match 17.8%; Score 28; DB 4; Length 41815;
Best Local Similarity 58.3%; Pred. No. 38;
Matches 49; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 62 AGCCTTGCTTATTTGAAACCAAGAGATACCTACCCAAACATCCATCTTACTCATGC 121
DB 33058 AGACATGAGATTCACACACATCCGCATAGGCTCTTACGATCCCTGGATCATTC 33117
QY 122 AACTCCATGCAACACGACATA 145
DB 33118 ATGTACATAGACATGATGTA 33141

RESULT 11
US-09-949-016-191201/c
;; Sequence 191201, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:


```
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 191201
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-191201
```

```
Query Match          17.7% Score 27.8; DB 4; Length 601;
Best Local Similarity 58.0%; Pred. No. 7.9;
Matches 47; Conservative 1; Mismatches 33; Indels 0; Gaps 0;
```

```
OY 57 GTCCAGCCTTGCCTATTCTGAAACCAAGAGATACCTACTCCCAACATCTTACT 116
DB 322 GTCCAGCCTTGCCTATTCTGAAACCAAGAGATACCTACTCCCAACATCTTACT 263
OY 117 CATGCACTTCATGCAACA 137
DB 262 CATGCAATAGCATGGAATA 242
```

```
RESULT 12
US-09-949-016-14149/c
Sequence 14149, Application US/09949016
Patent No. 6812339
```

```
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 14149
LENGTH: 11490
TYPE: DNA
ORGANISM: Human
US-09-949-016-14149
```

```
Query Match          17.7% Score 27.8; DB 4; Length 11490;
Best Local Similarity 52.1%; Pred. No. 26;
```

```
Matches 62; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
OY 38 CCGCTTCCTTGTTCATCTGCAAGCCTTGCTATTCTGAACCAAGAGATACCTACTC 97
DB 5642 CAGGGCCTCTGCTCCCTGCCCCCTGCCCCCACTCAGAGAGAGAGAGAGAGAGAG 583
OY 98 CCAACAAATCATCTTACTATGCACTTCATGCAACGACGACATATGTTCTGAA 156
DB 5582 CTTTGACACCTGATGCAATCCGCTTGAGTGAGGAGGAGGAGGAGGAGGAGGAGG 5524
```

```
RESULT 13
US-09-949-016-16740
```

```
Sequence 16740, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 16740
LENGTH: 14952
TYPE: DNA
ORGANISM: Human
US-09-949-016-16740
```

```
Query Match          17.7% Score 27.8; DB 4; Length 14952;
Best Local Similarity 59.5%; Pred. No. 29;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
OY 67 TGCCATTCTGAAACCAAGAGATACCTACTCCCAACATCTTACTCATGCAACTT 126
DB 70 TGTCAACCGTATGATCCAGATGATATTAATGTTAGGAATCTCTCTCAGAGTAAC 129
OY 127 CCATGCAACGCGACATA 145
DB 130 GGTTTACACACACACACA 148
```

```
RESULT 14
US-09-710-279-1813/c
Sequence 1813, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: P03480US
CURRENT APPLICATION NUMBER: US/09/710,279
PRIORITY FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: 60/164,258
PRIORITY FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1813
LENGTH: 1413
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: nucleic acid sequence
US-09-710-279-1813
```

```
Query Match          17.5% Score 27.4; DB 4; Length 1413;
Best Local Similarity 55.9%; Pred. No. 15;
Matches 52; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
```

```
OY 14 ATGCACTACTACATGCTTGTCAACCGTCTGTTGTCATGTCAGCCTTGCTTAT 73
DB 1213 ATTGAATTTTACGTGTGCTGTTCAAGATTTGTTTATCTTATCATATATATCATCATAC 1154
OY 74 TCTGAACCAAGAGATACCTACTCCCAACAT 106
DB 1153 CAAGATTCGAATATATCTTGTAACTCTCTAT 1121
```

```
RESULT 15
```



```

US-09-134-001C-1413/C
: Sequence 1413. Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR APPLICATION NUMBER: 1998-08-13
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 1413
: LENGTH: 1428
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1413

```

Query Match	17.5%	Score 27.4;	DB 3;	length 1428;
Best Local Similarity	55.9%;	Pred. No. 15;		
Matches	52;	Conservative	0;	Mismatches 41;
				Indels 0;
				Gaps 0;

Qy	Db	Qy	Db
14	1228	74	1168
ATGGCATCTCATCTCTGTTCAACGGTTCGCTTGTCATTCGTCGAAGCTTCCAT	ATTGAATCTACGTCGCTCGTCCACAGTTTGTTATCTTCATCTGATTAATTCATCCATAC	TCTGATCCGAAGATCTACTCTCCGAACAT	CAAGATGCAATATATCTTGTAATCTCTAT
73	1168	106	1136

Search completed: September 15, 2005, 08:25:04
Job time : 31.7481 secs


```

Db      15 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 74
Qy      61 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 120
Db      75 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 134
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 157
Db      135 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 171

```

RESULT 2

US-10-713-381-1

```

; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

```

```

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 60
Db      1155 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 1214
Qy      61 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 120
Db      1215 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 1274
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 1311

```

RESULT 3

US-10-713-381-2

```

; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

```

```

Query Match      100.0%; Score 157; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 3.1e-43;
Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 60
Db      1155 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 1214
Qy      61 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 120
Db      1215 AAGCCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 1274
Qy      121 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 157
Db      1275 CAACCTTCATGCAACACGACATATGTTTCTCGAAC 1311

```

RESULT 4

US-10-713-381-3

```

; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

```

```

Query Match      93.0%; Score 146; DB 20; Length 158;
Best Local Similarity 99.4%; Pred. No. 8.2e-40;
Matches 157; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

Qy      1 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 59
Db      1 CGGTGATCTCAACATGGCATACATGCTGTGTTCAACGGTTCGTTCCTTCATGTC 60
Qy      60 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 119
Db      61 CAAGCTTGCTTATTTCTGAACCAAGAGATACCTACTCCCAACATCCATCTTACTCAG 120
Qy      120 GCAACTTCATGCAACACGACATATGTTTCTCGAAC 157
Db      121 GCAACTTCATGCAACACGACATATGTTTCTCGAAC 158

```

RESULT 5

US-10-713-381-5

```

; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAAT, CARL W.

```



```

; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match          31.8%; Score 50; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 4,6e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      75 CTGAACCAAGAGTACTCTCCCAACATCCATCTTACTCATGCAAC 124
Db      1 CTGAACCAAGAGTACTCTCCCAACATCCATCTTACTCATGCAAC 50

RESULT 6
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNMAN, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match          25.5%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      85 AGGATCTACTCTCCCAACATCCATCTTACTCATGCAAC 124
Db      1 AGGATCTACTCTCCCAACATCCATCTTACTCATGCAAC 40

RESULT 7
US-10-674-124A-9081/C
; Sequence 9081, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideotoehi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
```

```

; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 9081
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: chr5.fa.07f.rz.139538206
; FEATURE:
; OTHER INFORMATION: Located on chromosome 5
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; FEATURE:
; OTHER INFORMATION: sequence : 128437212
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 250186
US-10-674-124A-9081

Query Match          20.5%; Score 32.2; DB 19; Length 399;
Best Local Similarity 61.2%; Pred. No. 1.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY      72 ATTCTGAACCAAGAGTACTCTCCCAACATCCATCTTACTCATGCAACTTCGACG 131
Db      239 ATTCTGAAGAAAACAGTGGGCCACCTCAAAAACCAACATCTTACTTAACTTAAAAA 180

QY      132 CAACACGACATATGTTTCCTGAA 156
Db      179 AAAAAAACACACATTTGTGATGAA 155

RESULT 8
US-10-282-122A-36272/C
; Sequence 36272, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
```



```
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO 36272
/ LENGTH: 2352
/ TYPE: DNA
/ ORGANISM: Streptococcus mutans
US-10-282-122A-36272
```

```
Query Match      20.4%; Score 31.6; DB 17; Length 2352;
Best Local Similarity 58.5%; Pred. No. 3.9;
Matches 55; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
```

```
QY 62 AGCTTGCCTATTGTGACCAAGAGTACTCTCCCAACAATCATCTTACTATGC 121
    |||||
DB 234 ATCTGACTGTGCCCAAGCAAGCAAGAGCATCTCTGAAAAATCTGTGCTACTATGC 175
    |||||
QY 122 AACTTCATGCAACAGCAGCATATGTTCCTGA 155
    |||||
DB 174 ATTTCAGCAAAAGAGCATTTTCTCTGA 141
    |||||
```

RESULT 9

```
US-10-437-963-1496/c
/ Sequence 1496, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 1496
/ LENGTH: 1122
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_101357C.1
US-10-437-963-1496
```

```
Query Match      19.6%; Score 30.8; DB 19; Length 1122;
Best Local Similarity 57.1%; Pred. No. 5.6;
Matches 56; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```
QY 60 CAAGCTTGCTATTGTGAACAAGATACCTACCTCCCAACAATCATCTTACTCAT 119
    |||||
DB 313 CAGCTCGCATATTGTGATCCCGGGATCATCTAAATCAACAAGCCATATTCTTCT 254
    |||||
QY 120 GCAACTTCATGCAACAGCAGCATATGTTCCTGAAC 157
    |||||
DB 253 TCAGATCCACAATTCGACCCCTCCCATATCTGCTGAAC 216
    |||||
```

```
RESULT 10
US-10-027-632-244811
```

```
/ Sequence 244811, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 244811
/ LENGTH: 559
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-244811
```

```
Query Match      19.5%; Score 30.6; DB 13; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 24 ACATGCTGTTCACCGTTCCTTGTTCATGTGTCACAGCCTGCTATTCTGACCA 83
    |||||
DB 380 ACATGCTATTTCAGCTTTCGATTTGCTCATCTTGCGAGGTACTTTTCTAGATTCA 439
    |||||
QY 84 GAGGATTAATCTCTCCCAACAATCATCTTACTCATGGAATTCATGCAACAGCACA 143
    |||||
DB 440 TCTGCTCCCTACTTTAAAGATTCATGGAAGTCTCAAAATCCATAGCAGCACTGAAC 499
    |||||
QY 144 TATGTTCTCTGAA 156
    |||||
DB 500 ATTTCCTCCCA 512
    |||||
```

RESULT 11

```
US-10-027-632-244811
/ Sequence 244811, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827,129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
```


NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 244811
LENGTH: 559
TYPE: DNA
ORGANISM: Human
US-10-027-632-244811

Query Match 19.5%; Score 30.6; DB 17; Length 559;
Best Local Similarity 51.9%; Pred. No. 5.1;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 24 ACATGCTGTTCACCGTGTGCTTGTTCATGCGCAAGCGTCCATTCTGAACCA 83
DB 380 ACATGCTGTTCACCGTGTGCTTGTTCATGCGCAAGCGTCCATTCTGAACCA 439
QY 84 GAGGATACCTTACCTCCAAACATCCATCTTACTGATGCAACTTCATGCAACAGCA 143
DB 440 TCTGCTCCCTTAAAGATTCATGGAAGTCTTAAATCCATGACACTGAC 499
QY 144 TATGTTCTCTGAA 156
DB 500 ATTTCCTCCCA 512

RESULT 12

US-10-052-482-223
Sequence 223, Application US/10052482
Publication No. US20040072268A1
GENERAL INFORMATION:
APPLICANT: Morriss, David
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
FILE REFERENCE: A-71087/RMS/DCF
CURRENT APPLICATION NUMBER: US/10/052,482
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 241
SOFTWARE: PatentIn version 3.1
SEQ ID NO 223
LENGTH: 79467
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5502)..(5521)
OTHER INFORMATION: "n" at positions 5502 to 5521 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4095)..(4369)
OTHER INFORMATION: "n" at positions 4095 to 4369 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11089)..(10620)
OTHER INFORMATION: "n" at positions 10089 to 10620 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13273)..(13370)
OTHER INFORMATION: "n" at positions 13273 to 13370 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (20762)..(20781)
OTHER INFORMATION: "n" at positions 20762 to 20781 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (30916)..(30916)
OTHER INFORMATION: "n" at positions 30751 to 30916 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (46579)..(46772)

OTHER INFORMATION: "n" at positions 46759 to 46772 can be any base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (77382)..(77401)
OTHER INFORMATION: "n" at positions 77372 to 77401 can be any base
US-10-052-482-223

Query Match 19.5%; Score 30.6; DB 18; Length 79467;
Best Local Similarity 56.4%; Pred. No. 31;
Matches 57; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 51 TCCATGCTCCAGCGTGTGCTTGTTCGACCAAGAGATCTACTCCCAATTCAT 110
DB 53779 TCCATGCTCCAGCGTGTGCTTGTTCGACCAAGAGATCTACTCCCAATTCAT 53838
QY 111 CTTCATGCACTTCATGCAACAGCAACATATGTTTC 151
DB 53839 CACACACACACACACACACACACACACAGCTTGCGTC 53879

RESULT 13

US-10-317-273-11
Sequence 11, Application US/10317273
Publication No. US20040110158A1
GENERAL INFORMATION:
APPLICANT: Kenneth W. Dobie
TITLE OF INVENTION: MODULATION OF MITOCHONDRIAL RIBOSOMAL PROTEIN L13 EXPRESSION
FILE REFERENCE: RTS-0478
CURRENT APPLICATION NUMBER: US/10/317,273
CURRENT FILING DATE: 2002-12-10
NUMBER OF SEQ ID NOS: 159
SEQ ID NO 11
LENGTH: 52001
TYPE: DNA
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11121, 11122, 11123, 11124, 11125, 11126, 11127, 11128, 11129,
LOCATION: 11130, 11131, 11132, 11133, 11134, 11135, 11136, 11137,
LOCATION: 11138, 11139, 11140, 11141, 11142, 11143, 11144, 11145,
LOCATION: 11146, 11147, 11148, 11149, 11150, 11151, 11152, 11153
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11154, 11155, 11156, 11157, 11158, 11159, 11160, 11161, 11162,
LOCATION: 11163, 11164, 11165, 11166, 11167, 11168, 11169, 11170,
LOCATION: 11171, 11172, 11173, 11174, 11175, 11176, 11177, 11178,
LOCATION: 11179, 11180, 11181, 11182, 11183, 11184, 11185, 11186
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11187, 11188, 11189, 11190, 11191, 11192, 11193, 11194, 11195,
LOCATION: 11196, 11197, 11198, 11199, 11200, 11201, 11202, 11203,
LOCATION: 11204, 11205, 11206, 11207, 11208, 11209, 11210, 11211,
LOCATION: 11212, 11213, 11214, 11215, 11216, 11217, 11218, 11219
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc_feature
LOCATION: 11220
OTHER INFORMATION: n = A,T,C or G
US-10-317-273-11

Query Match 19.4%; Score 30.4; DB 19; Length 52001;
Best Local Similarity 61.2%; Pred. No. 31;
Matches 49; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 50 TTTCATGTCACAGCGTGTGCTTGTGACCAAGAGATCTACTCCCAATTCAT 109
DB 20055 TTTCATGTCACAGCGTGTGCTTGTGACCAAGAGATCTACTCCCAATTCAT 20114
QY 110 TCTTACTCAGCACTTCCA 129

Db 20115 TCTTCTCAGCAAACTTCTA 20134

RESULT 14

US-10-713-381-4
; Sequence 4, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFEMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-4

Query Match 19.1%; Score 30; DB 20; Length 30;

Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 CATGCTGTTCACCGTTCGTTGTTCCA 54
Db 1 CATGCTGTTCACCGTTCGTTGTTCCA 30

RESULT 15

US-10-027-632-170626
; Sequence 170626, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/195,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 170626
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-170626

Query Match 19.1%; Score 30; DB 13; Length 843;
Best Local Similarity 51.5%; Pred. No. 9.5;
Matches 69; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 6 CATTCACATGGCATACTACATCTTGTTCACCGTTCGTTGTTCCATGTCACAGCC 65
Db 552 CAGCACCCAGGCCAAGCTCAAGACAAAGTCAAGCTCTCTGTTGGGGCCACCATTCATGCC 611
Qy 66 TTGCCTATTCTGAACCAAGAGATACCTATCCAAACATGATCTTATCTGATGCAACT 125
Db 612 TTGCATGTGCTGGAAGATGGGGGCTTCGCTCCTGCTCATCAACGCTGGGAGGCGCTCC 671
Qy 126 TCCATGCAAAACACG 139
Db 672 TGCATCCGAACACG 685

Search completed: September 15, 2005, 20:45:38
Job time : 108.248 secs

Db 421 C G T G T C A T C T C A C A T G G C A T A C T A C A T G C T T G T T C A A C C G T T C G T C T T G T T C C A T C G T C C 362

QY 61 AACCTTGGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 120
 DB 361 AAGCTTGCTGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 302

QY 121 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 157
 DB 301 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 265

RESULT 2
 CG224225 915 bp DNA linear GSS 22-AUG-2003
 LOCUS OGIAG08TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0716B15,
 DEFINITION genomic survey sequence.
 ACCESSION CG224225 GI:34124113
 VERSION CG224225
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 915)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 JOURNAL Other GSSs: OGIAG08TH
 COMMENT Contact: Cathy WhiteIaw
 TIGR

7912 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteIaw@tigr.org
 Seq primer: TP
 Classes: sheared ends.

FEATURES
 source Location/Qualifiers
 1..915
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0716B15"
 /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 100.0%; Score 157; DB 9; Length 915;
 Best Local Similarity 100.0%; Pred. No. 1.9e-39;
 Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCACTCAGATGGCATCTACTACATGCTTGTTCACCGCTTGCTTCCATCGTCC 60
 DB 534 CGTGTCACTCAGATGGCATCTACTACATGCTTGTTCACCGCTTGCTTCCATCGTCC 593

QY 61 AAGCTTGGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 120
 DB 594 AAGCTTGGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 653

QY 121 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 157
 DB 654 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 690

RESULT 3
 CC656933 963 bp DNA linear GSS 19-JUN-2003
 LOCUS OGMDD20TM ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0554D15,
 DEFINITION genomic survey sequence.
 ACCESSION CC656933

VERSION CC656933.1 GI:32060225
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 963)
 WhiteIaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 JOURNAL Other GSSs: OGMDD20TV
 COMMENT Contact: Cathy WhiteIaw
 TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whiteIaw@tigr.org
 Seq primer: TP
 Classes: sheared ends.

FEATURES
 source Location/Qualifiers
 1..963
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA0554D15"
 /note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN

Query Match 100.0%; Score 157; DB 9; Length 963;
 Best Local Similarity 100.0%; Pred. No. 2e-39;
 Matches 157; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGTGTCACTCAGATGGCATCTACTACATGCTTGTTCACCGCTTGCTTCCATCGTCC 60
 DB 752 CGTGTCACTCAGATGGCATCTACTACATGCTTGTTCACCGCTTGCTTCCATCGTCC 811

QY 61 AAGCTTGGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 120
 DB 812 AAGCTTGGCTATTTCGAACCAAGAGATCTACTCTCCAAACATCCATTCTTACTCATG 871

QY 121 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 157
 DB 872 CAACCTTCATGCAAAACGCGACATATGTTCTCTGAAC 908

RESULT 4
 CE588104/c 636 bp DNA linear GSS 28-SEP-2003
 LOCUS CE588104/c
 DEFINITION tigr-gss-dog-17000366359543 dog library Canis familiaris genomic,
 genomic survey sequence.
 ACCESSION CE588104
 VERSION CE588104.1 GI:36904885
 KEYWORDS GSS.
 SOURCE Canis familiaris (dog)
 ORGANISM Canis familiaris
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 1 (bases 1 to 636)
 Kirkness,B.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
 Ruesch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
 Venter,J.C.
 The dog genome: survey sequencing and comparative analysis
 Science 301 (5641), 1898-1903 (2003)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekiknes@tigr.org
Class: shotgun.

FEATURES

source

Location/Qualifiers
1..636
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 24.1%; Score 37.8; DB 9; Length 636;
Best Local Similarity 58.4%; Pred. No. 0.39; Mismatches 47; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 TCTTGTTCATCGTCGACGCTTGCTATCTGGAACGAGATACCTACCTCCCAACA 104
DB 292 TATTTTCCAGTTCAGGATCTGTTGGACCCAGGATATACCTCATTTTAAAA 233
QY 105 ATCCATCTTACTGATGCACTTCATGCAACGACGACATATGTTCTGAC 157
DB 232 ATACATTTTATTTATGAAATTTTCMAACACAAAATTAATGTTAGTGAAC 180

RESULT 5
LOCUS CO689495 643 bp mRNA linear EST 26-JUL-2004
DEFINITION DG11-25h6 DG11-kidney Canis familiaris cDNA 3', mRNA sequence.
ACCESSION CO689495
VERSION CO689495.1 GI:50638161
KEYWORDS EST.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 643)
Schlueter, T., Hermann, J., Weindel, M., Schuette, D., Kranz, H.,
Henrich, J., and Loeber, R.
Dog arraytag cDNA clone collection
Unpublished (2004)
Contact: Thomas Schlueter
LION Bioscience AG
Walhoferstrasse 98, D-69123 Heidelberg, Germany
Tel: +49 6221 4038 150
Fax: +49 6221 4038 290
Email: Thomas.Schueter@lionbioscience.com.
Location/Qualifiers

FEATURES

source

1..643
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/clone_lib="Kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="DG11-kidney"
/note="Organ: Kidney; Vector: Dog pBlueScript LION"

ORIGIN

Query Match 24.1%; Score 37.8; DB 7; Length 643;
Best Local Similarity 58.4%; Pred. No. 0.4; Mismatches 47; Indels 0; Gaps 0;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 45 TCTTGTTCATCGTCGACGCTTGCTATCTGGAACGAGATACCTACCTCCCAACA 104
DB 333 TATTTTCCAGTTCAGGATCTGTTGGACCCAGGATATACCTCATTTTAAAA 274

QY 105 ATCCATCTTACTGATGCACTTCATGCAACGACGACATATGTTCTGAC 157
DB 273 ATACATTTTATTTATGAAATTTTCMAACACAAAATTAATGTTAGTGAAC 221

RESULT 6
LOCUS AG366773 1085 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-170B12.TJ, genomic survey sequence.
ACCESSION AG366773
VERSION AG366773.1 GI:47977978
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1085)
Hattori, M., Toyoda, A., Noguchi, H., Kojima, T. and Sakaki, Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC library availability, please contact Kuniya Abe (abe@rc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9159
e-mail: abe@rc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : DBAC3.6
Vector : BclRI
R.site 1 : BclRI
R.site 2 : BclRI.
Location/Qualifiers

COMMENT

1..1085
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-170B12.TJ"
/sex="male"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 24.1%; Score 37.8; DB 9; Length 1085;
Best Local Similarity 51.9%; Pred. No. 0.46; Mismatches 64; Indels 0; Gaps 0;
Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 22 CTACATGCTTGTTCACGCTGCTGCTTGTTCATGTCGACGCTGCTATCTGAC 81
DB 719 CTCCCT 660
QY 82 AAGAGATACCTACTCCCAACGATCTTACTGATGCAACCTTCATGCAACGCA 141
DB 659 ACATNATATACNCCGCCNANNANAAACANNANAAACCCGCAATTTTANANANATAAA 600

QY 142 CATATGTTCCG 154
DB 599 CAAACTTTTTTG 587

RESULT 7
BS505306

LOCUS BB505306 700 bp mRNA linear EST 25-OCT-2001
DEFINITION BB505306 RIKEN full-length enriched, 10 days lactation, adult female mammary gland Mus musculus cDNA clone D730002G06, mRNA sequence.
ACCESSION BB505306
VERSION BB505306.2 GI:16442791
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 700)
REFERENCE 1 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
TITLE JOURNAL
COMMENT On Jul 27, 2000 this sequence version replaced gi:9514268.
Contact: Yoshinide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kikura, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES
source
1. .700
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D730002G06"
/sex="female"
/tissue_type="mammary gland"
/dev_stage="10 days lactation, adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days lactation, adult female mammary gland"
/note="Site_1: SalI; Site_2: BamHI, cDNA library was prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCCGCACTCGAGTGTGTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCTCGAGTTATTAATTAATCCCGCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FHC I."

ORIGIN

Query Match 21.7%; Score 34; DB 2; Length 700;
Best Local Similarity 54.9%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 27 TCGTGTCAACCGTTCGTCCTGTCATCGCCAAAGCCTTGCTTAACCAAGAG 86
DB 195 TTTCTTTTTCCTCCCTCCCTCTGATTTCTTTGACTTCACATTTTGCCCTTAAGTT 254
QY 87 GATACCTACTCCCAACAATCCATCTTAATGCAACTTCATGCAACGACATAT 146
DB 255 GTTCTAGATTCCTCAAGAAAGCATTTTGACATGACATGTAACATATCAACCTGAAAT 314
QY 147 GT 148
DB 315 GT 316

RESULT 8

LOCUS CL026582 1677 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-23G24 Sp6.1 CH216 Xenopus tropicalis genomic clone
CL026582
CL026582.1 GI:40470443

ACCESSION CL026582
VERSION CL026582.1 GI:40470443
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Silurana.

REFERENCE 1 (bases 1 to 1677)
AUTHORS Krenitschki, C., Carter, J., McPherson, J., Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends
High quality sequence start: 125
High quality sequence stop: 217.
Location/Qualifiers

FEATURES

source
1. .1677
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-23G24"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"

ORIGIN

Query Match	21.4%	Score 33.6	DB 9	Length 1677
Best Local Similarity	53.9%	Pred. No. 12		
Matches 69	Conservative 0	Mismatches 59	Indels 0	Gaps 0
Db	3	TGTCATCTCACATGGCAGTACTACATGCTGTGTTCAACCGCTGCTGTGTTCCATGCTCAAC	62	
Qy	769	TGTGCTCTCACATGGCAGTACTACATGCTGTGTTCAACCGCTGCTGTGTTCCATGCTCAAC	828	
Db	829	TCATCAACATATCTTACAGATTCATGCTCAACCTCTCACTTCTTATATATTATTCAGACA	888	
Qy	123	ACTTCAT 130		
Db	889	TCTTCAT 896		
RESULT 9				
LOCUS	A1744861	440 bp	mRNA	linear EST 21-JUN-1999
DEFINITION	tr16b06.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2218451 3'			
ACCESSION	A1744861			
VERSION	A1744861.1	GI:51133149		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 440)			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapdb-remail.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLND at: www.bio.liml.gov/bdtp/image/image.html Seq primer: -40UP from Gibco High quality sequence stop: 411. Location/Qualifiers			
FEATURES				
SOURCE	1..440			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="IMAGE:2218451"			
	/tissue_type="tumor, 5 pooled (see description)"			
	/lab_host="DH10B"			
	/clone_lib="NCI_CGAP_Ov23"			
	/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"			
ORIGIN				
Query Match	20.9%	Score 32.8	DB 1	Length 440
Best Local Similarity	56.5%	Pred. No. 14		
Matches 61	Conservative 0	Mismatches 47	Indels 0	Gaps 0
Db	46	CTGTGTCATGCTCCAGCCTTGCTGCTATTCTGAAACCAAGAGATACCTACTCCAAACA	105	
Qy	259	CTTGTGCGACGAAATAGATATATCTATCTCTCCCAAGCAACCTCTACTTATTT	318	
Db	106	TTCATCTTACATGCACTTCCATGCAACGACGACATATGTTCT	153	

DB	319	TCGATCTTAATTAAGACCGTCGATGTCGCTAGATATCTGTTCCCT	366
RESULT 10			
CR575123			
LOCUS			
DEFINITION	CR575123 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA024m23		
ACCESSION	CR575123		
VERSION	CR575123.1		
KEYWORDS	EST.		
SOURCE	Xenopus tropicalis (western clawed frog)		
ORGANISM	Xenopus tropicalis		
REFERENCE			
AUTHORS	Croning, M.D.R., Ashurst, J.L., Taylor, R., Garrett, N. and Rogers, J.		
TITLE	Sanger Xenopus tropicalis EST project 2001 (2004)		
JOURNAL	Unpublished (2004)		
COMMENT	Contact: Croning MDR Sanger Institute Hinxton, Cambridgeshire, CB10 1SA, UK Email: trop@sanger.ac.uk Sanger Xenopus tropicalis EST project 2001 TROPICALIS_SEQUENCE_ID: THdA024m23.p1bSP6 This sequence is from a Xenopus Gene Collection (XGC) library constructed by Nigel Garrett. Seq primer: SP6.		
FEATURES			
source	Location/Qualifiers 1..804 /organism="Xenopus tropicalis" /mol_type="mRNA" /db_xref="taxon:8364" /clone="THdA024m23" /dev_stage="tailbud head (stage 28-30)" /lab_host="Escherichia coli DH10B." /clone_id="XGC-tailbud-head" /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from tailbud head. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."		
ORIGIN			
Query Match	20.8%; Score 32.6; DB 7; Length 804;		
Best Local Similarity	63.3%; Pred. No. 20;		
Matches	50; Conservative 0; Mismatches 29; Indels 0; Gaps 0;		
QY	74 TCTGAACCAAGAGATACCTAGCTCCCAACATCATCTTACTGATGCAACTTCATGCA 133		
DB	480 TCTGAATTTATGAAAGAGCATCTCCATCCATGACTCATTTATGCAAGTAATTTACATTTT 539		
QY	134 AACAGCACAATAGTTTCC 152		
DB	540 AAAAAGTATTTTTC 558		
RESULT 11			
CNS003FV/c			
LOCUS			
DEFINITION	CNS003FV 1101 bp DNA linear GSS 03-JUN-1999		
ACCESSION	Drosophila melanogaster genome survey sequence TET3 end of BAC #		
VERSION	BACR08003 of Rpci-98 library from Drosophila melanogaster (fruit		
KEYWORDS	fly), genomic survey sequence.		
SOURCE	AL064657		
ORGANISM	AL064657.1 GI:4942009		
REFERENCE	GSS.		
AUTHORS	Drosophila melanogaster (fruit fly)		
COMMENT	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope.		


```

TITLE      Direct Submission
JOURNAL    Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : sege@genoscope.cns.fr
COMMENT     - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutyo Osoegawa and
            Aaron Mamosser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES   Location/Qualifiers
            source             1..1101
                                /organism="Drosophila melanogaster"
                                /mol_type="genomic DNA"
                                /db_xref="taxon:7227"
                                /clone="BACR08003"
                                /clone_11b="RPCI-98"
                                /note="end : TET3"

ORIGIN
Query Match      20.8%; Score 32.6; DB 9; Length 1101,
Best Local Similarity 21.9%; Pred No. 22;
Matches          25; Conservative 55; Mismatches 34; Indels 0; Gaps 0.

Dy      29  CTGTTCACCGCTGCTGCTTGTTCATCGTCACAGCTTGCTATTCTGAACAGAGA 88
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      997  CBNWTKDAAYBDKXDATYKBSKYCCMAAVSCCARNDNTWAMCHDKKYMCHWHAAHAA 938
          ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Dy      89  TACTACTCTCCCAACATCATCTTACTGATGCAACTTTCATGCAACACGCAC 142
          ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      937  WDNBKKKBCBMAAMABRCRDKDKTYYKXTKBCYCYKAKAGRSYKSCC 884

RESULT 12
LOCUS      BE690384                               347 bp      mRNA      linear      EST 11-SEP-2000
DEFINITION uw66f12.y1 Soares mammary gland NMLMG Mus musculus cDNA clone
IMAGE:3466991 5', mRNA sequence.
ACCESSION  BE690384
VERSION    BE690384.1 GI:10078008
KEYWORDS   EST.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 347)
            NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgsb@r-rtmail.nih.gov
            This clone is available royalty-free through LINT; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1386351
FEATURES   Seg primer: -40RP from Gibco
            High quality sequence scop: 270.
            Location/Qualifiers
            source             1..347
                                /organism="Mus musculus"
                                /mol_type="mRNA"
                                /db_xref="taxon:10090"
                                /clone="IMAGE:3466991"
                                /sex="female (lactating)"

```

			/tissue_type="mammary gland"	
			//lab_host="DH10B"	
			//clone_id="Soares_mammary_gland_NMLMG"	
			/note="Vector: pT73D-Pac (Pharmacia) with a modified	
			polylinker; 1st strand cDNA was prepared from mammary	
			gland tissue from a lactating female, and was then primed	
			with a Not I - oligo(dT) primer. Double-stranded cDNA was	
			ligated to Eco RI adaptors (Pharmacia), digested with Not	
			I and cloned into the Not I and Eco RI sites of the	
			modified pT73 vector. Library is normalized. Library	
			was constructed by Bento Soares and M. Fatima Bonaldo."	
ORIGIN				
Query Match	20.6%	Score 32.4;	DB 2;	Length 347;
Best Local Similarity	54.1%	Pred. No. 18;		
Matches	66;	Conservative	0;	Mismatches 56; Indels 0; Gaps 0;
Dy	27	TGCTTTTCACACCGTGCGTCTGTTCATGCATGTCACAACCCCTTGCTATTGCAACCAAG	86	
Dd	148	TTCTTTTTCACCTCCCTCCTTGATTTCTTTGGACTTCACCTTTCATTTTGGCCTTTAAGTT	207	
Dy	87	GATACCTACTGCCAACATCATCTTACTCATGCACTTCACATGCAACAGCACATAT	146	
Dd	208	GTTCTAGATTCCCAAAGACCATTTTGCATACAGTCAACATACATACACCTGAAT	267	
Dy	147 GT 148			
Dd	268 GT 269			
RESULT 13				
AJ746802/c				
LOCUS				
DEFINITION	AJ746802 forward - stimulated minus unstimulated macrophage Sus	483 bp mRNA linear EST 07-JUN-2004		
ACCSSION	AJ746802	scrofa cDNA clone F_00001825C_D07, mRNA sequence.		
VERSION	AJ746802			
KEYWORDS	AJ746802.1 GI:49916882			
SOURCE	EST.			
ORGANISM	Sus scrofa (pig)			
COMMENT	Sus scrofa			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.			
TITLE	1 (bases 1 to 483)			
JOURNAL	Hopwood,P.A., Zhang,F., Lowden,S., Talbot,R., Burt,D., Archibald,A.			
COMMENT	Development of a porcine cDNA microarray unpublished (2004) Contact: Hopwood PA Dept. of Preclinical Veterinary Sciences Royal School for Veterinary Studies Summerhall, Edinburgh, EH9 1QH, UNITED KINGDOM Sequencing was performed by ARK genomics. This clone is available from ARK- Genomics, Roslin Institute, Roslin, Midlothian EH25 9PS, UK. See www.ark-genomics.org or contact info@arkgenomics.org. Location/Qualifiers 1..483 /organism="Sus scrofa" /mol_type="mRNA" /db_xref="taxon:9823" /clone="F_00001825C_D07" /tissue_type="lung" /cell_type="macrophage" /clone_id="forward - stimulated minus unstimulated macrophage"			
FEATURES				
source				
ORIGIN				
Query Match	20.6%	Score 32.4;	DB 1;	Length 483;
Best Local Similarity	55.3%	Pred. No. 20;		
Matches	63;	Conservative	0;	Mismatches 51; Indels 0; Gaps 0;
Dy	21	ACTACATGCTTTTCACACCGTGTGCTCTTGTTCATGTCATGTCACAACCGTTGCTATTTCGAC	80	
Dd	371	ACGACATCTTAGTAAGAAGCTTTTTTTTTTTTTAAAAAAAGCCCAATATATTTCTTAAC	312	

QY 81 CAGAGGATGCTACTCTCCCAAGATTCATCTACTGATGCACTTCGATGCA 134
DB 311 ACAGTGAAGTCCAGCTACCTATCATTCAGTTAAATTAAGAAATTTCCAGGAA 258

RESULT 14
LOCUS AW990724/c 497 bp mRNA linear EST 02-JUN-2000
DEFINITION ufl0b11.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1510941 3', mRNA sequence.

ACCESSION AW990724
VERSION AW990724
KEYWORDS AW990724.1 GI:8186451
SOURCE EST.
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 497)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: ufl0b11.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:937793
High quality sequence stop: 459.
Location/Qualifiers
1..497
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:1510941"
/sex="female (lactating)"
/issue_type="mammary gland"
/lab_host="DH10B"
/lab_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 20.6%; Score 32.4; DB 2; Length 497;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TGCTGTGAACGCTGCTGTGTCATGTCGCAAGCCTTGCTATTGCAACCAAG 86
DB 490 TTTCTTTTTCACCTCTCTGTGATTTCTTGTACTTCACCTTATTTGGCCTTTAAGTT 431

QY 87 GATACCTACTCCCAACATCATCTTACTGATGCAACTTCATGCAACAGCAGATAT 146
DB 430 GTTCTAGATTCCCAAGAGCCATTTTGACCTAGCACTAAGCAATACACCTGAAT 371

QY 147 GT 148
DB 370 GT 369

RESULT 15
BE632212/c 498 bp mRNA linear EST 25-AUG-2000
LOCUS BB632212
DEFINITION uul1f01.x1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:3371641 3', mRNA sequence.

ACCESSION BE632212

VERSION BE632212.1 GI:9914830
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 498)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other ESTs: uul1f01.y2
Contact: Robert Strausberg, Ph.D.
Email: cgapdb-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1081245
High quality sequence stop: 459.
Location/Qualifiers
1..498
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3371641"
/sex="female (lactating)"
/issue_type="mammary gland"
/lab_host="DH10B"
/lab_lib="Soares_mammary_gland_NMLMG"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 20.6%; Score 32.4; DB 2; Length 498;
Best Local Similarity 54.1%; Pred. No. 20;
Matches 66; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 27 TGCTGTGAACGCTGCTGTGTCATGTCGCAAGCCTTGCTATTGCAACCAAG 86
DB 491 TTTCTTTTTCACCTCTCTGTGATTTCTTGTACTTCACCTTATTTGGCCTTTAAGTT 432

QY 87 GATACCTACTCCCAACATCATCTTACTGATGCAACTTCATGCAACAGCAGATAT 146
DB 431 GTTCTAGATTCCCAAGAGCCATTTTGACCTAGCACTAAGCAATACACCTGAAT 372

QY 147 GT 148
DB 371 GT 370

Search completed: September 16, 2005, 08:08:25
Job time : 496.661 secs

This Page Blank (uspto)

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

RESULT 3
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX224394
Sequence 1 from Patent WO0160997.
AX224394
AX224394.1 GI:15554636
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 4
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX224395
Sequence 2 from Patent WO0160997.
AX224395
AX224395.1 GI:15554637
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;

REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 5
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD062176
Male tissue-preferred regulatory region and method of using same.
BD062176
BD062176.1 GI:22607781
JP 2001520523-A/1.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1394)
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and
Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN,TIMOTHY W FOX,CARL W GARNAAAT,GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82,C12N15/29,C12N9/24,C12N9/22,C12N9/10,C12N9/00 PC
,C07K14/34,C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH key Location/Qualifiers
1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 6
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD062177
Male tissue-preferred regulatory region and method of using same.
BD062177
BD062177.1 GI:22607782
JP 2001520523-A/2.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1394)
AUTHORS Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G.A. and Kendall,T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
COMMENT PN JP 2001520523-A/2
PD 30-OCT-2001
PF 19-JUN-1998 JP 1999504910
PR 23-JUN-1997 US 08/880499
PI MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAA, GARY A HUFFMAN,
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
PC C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source 1..1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTTCGCTGTTGCCA 30
1179 CATGCTGTTCACCGTTCGCTGTTGCCA 1208
Db
RESULT 7 3343 bp DNA linear PLN 12-MAY-2001
LOCUS AF360356
DEFINITION Zea mays male fertility protein (Me45) gene, complete cds.
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
Cloning of Me45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox,T.W., Trimnell,M.R. and Albertsen,M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
Location/Qualifiers
FEATURES
source 1..3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
/gene="Me45"
join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="Me45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Me45"
/codon_start=1
/product="male fertility protein"
/protein_id="AK52489.1"

/db_xref="GI:14028757"
/translation="MEKRLQMRRGDGI VOYPHLFFALALALLVADPGSLPLAEV
DTRPVHETLAPGEWGSMPDRNASLRRGRLEFGEVGPESIEFDLQGRPYGLA
DRLVVRMGEBAEWETPFAVNPDPMSSEVCANVSTTRQHKSEFCGRPLGRFGE
TSELVYADAYYGLMVGSGGVASSVAREADDPTRFANDLVHRNGSVFTDTSNR
SRKDLNLTLLBEGTGRLRLRYDPETSGVAVLKGIVFPNGVOISDHQFLPSETNG
RLMRVWLBGPRAGEVFPANLPGFPDNPVNSNGRQFWAIDCCRPDQGVFAPKRPMLR
TLVFKPPLSLKVLTVKARRMFTVLLALDDEGRVVEVLEDRGHEVWKLVSREVRGR
LWIGTVAHNHIAITPYPLED"
ORIGIN
Query Match 100.0%; Score 30; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTTCGCTGTTGCCA 30
1179 CATGCTGTTCACCGTTCGCTGTTGCCA 1208
Db
RESULT 8 158 bp DNA linear PAT 10-SEP-2001
LOCUS AX224396
DEFINITION Sequence 3 from Patent WO0160997.
ACCESSION AX224396
VERSION AX224396.1 GI:15554638
KEYWORDS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
FEATURES
source 1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"
ORIGIN
Query Match 80.7%; Score 24.2; DB 6; Length 158;
Best Local Similarity 89.7%; Pred. No. 11;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTTCGCTGTTGCC 29
25 CATGCTGTTCACCGTTCGCTGTTGCC 53
Db
RESULT 9 1304 bp DNA linear PAT 24-OCT-2002
LOCUS AX523502
DEFINITION Sequence 90 from Patent WO02064788.
ACCESSION AX523502
VERSION AX523502.1 GI:24412398
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1
Macina,R.A., Recipon,H., Chen,S.Y., Sun,Y. and Liu,C.
Compositions and methods relating to lung specific genes and
proteins
Patent: WO 02064788-A 90 22-AUG-2002;
Diadexus, Inc. (US)
Location/Qualifiers
FEATURES
source 1..1304
/organism="Homo sapiens"

ORIGIN /mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 73.3%; Score 22; DB 6; Length 1304;
Best Local Similarity 83.3%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATGCTGTTCAACCGTTCGTTGTTCCA 30
1155 CTTCTGTTCCCGCCGTTCTCTGTTCCA 1126

RESULT 10 AC124638 187192 bp DNA linear HTG 05-MAR-2004
AC124638/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1 map 17, *** SEQUENCING
LOCUS IN PROGRESS ***; 10 unordered pieces.

ACCESSION AC124638 GI:45120281
VERSION HTG; HTGS PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS MUS MUSCULUS (house mouse)
SOURCE MUS MUSCULUS
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 187192)
AUTHORS Birren, B., Nussbaum, C. and Lander, E.
TITLE Mus musculus chromosome 17, clone RP23-461H1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 187192)
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
JOURNAL Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L.,
AUTHORS Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,
AUTHORS Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
AUTHORS Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
AUTHORS Fero, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
AUTHORS Galagan, J., Gardyna, S., Ginde, S., Gold, S., Goyette, M., Graham, L.,
AUTHORS Grand-Pierre, N., Hage, B., Horton, L., Hulme, W., Iliev, I.,
AUTHORS Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., LaRocque, K.,
AUTHORS Lamasares, R., Landers, T., Lechoczy, J., Levine, R., Lindblad-Toh, K.,
AUTHORS Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N.,
AUTHORS Mathews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrum, J.,
AUTHORS Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
AUTHORS Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P.,
AUTHORS O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N.,
AUTHORS Pollara, V., Raymond, C., Retra, R., Rieback, M., Riley, R., Rise, C.,
AUTHORS Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
AUTHORS Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
AUTHORS Stojanovic, N., Strauss, N., Subramanian, A., Talmas, J., Tesfaye, S.,
AUTHORS Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,
AUTHORS Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
AUTHORS Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

REFERENCE 3 (bases 1 to 187192)
AUTHORS Direct Submission
TITLE Submitted (15-JUN-2002) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 187192)
AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
JOURNAL Anderson, M., Arachchi, H. M., Barina, N., Bastien, V., Bloom, T.,
AUTHORS Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y.,
AUTHORS Collymore, A., Cooke, P., Corum, B., Deatellano, K.,
AUTHORS Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
AUTHORS Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
AUTHORS Graham, L., Grand-Pierre, N., Hage, B., Horton, L., Hulme, W., Iliev, I.,
AUTHORS Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
AUTHORS Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
AUTHORS Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
AUTHORS MacDonald, P., Major, J., Manning, J., Mathews, C., McCarthy, M.,
AUTHORS Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
AUTHORS Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
AUTHORS O'Neill, D., Oliver, J., Peterson, K., Phunhngang, P., Pierre, N.,
AUTHORS Rachupka, A., Ramasamy, U., Raymond, C., Retra, R., Rise, C., Rogov, P.,
AUTHORS Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
AUTHORS Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

TITLE Talmas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
JOURNAL Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
AUTHORS Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
AUTHORS Direct Submission
TITLE Submitted (05-MAR-2004) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS On Mar 5, 2004 this sequence version replaced gi:28893670.
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: L24115
Center clone name: 461_H_1

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 9927 9926: contig of 9926 bp in length
* 9927 10026: gap of 100 bp
* 10027 20081: contig of 10055 bp in length
* 20082 20181: gap of 100 bp
* 20182 43259: contig of 23078 bp in length
* 43260 43359: gap of 100 bp
* 43360 45947: contig of 2586 bp in length
* 45948 46047: gap of 100 bp
* 46048 100940: contig of 54893 bp in length
* 100941 101040: gap of 100 bp
* 101041 106719: contig of 5679 bp in length
* 106720 106819: gap of 100 bp
* 106820 140370: contig of 33551 bp in length
* 140371 140470: gap of 100 bp
* 140471 147705: contig of 7235 bp in length
* 147706 147805: gap of 100 bp
* 147806 152565: contig of 4760 bp in length
* 152566 152665: gap of 100 bp
* 152666 187192: contig of 34527 bp in length.

FEATURES
source 1.187192
Location/Qualifiers

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/map="17"
/clone="RP23-461H1"
/clone_1ib="RP23-461H1"
/clone_1ib="RP23-461H1" Female Mouse BAC"

ORIGIN

Query Match 72.0%; Score 21.6; DB 2; Length 187192;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ATGCTTGTTCACCGTTCGTTGTTCC 29
52881 ATGCTTGTTCACCGTTCGTTGTTCC 52854

RESULT 11 AC151282 188427 bp DNA linear HTG 24-AUG-2004
AC151282/c MUS MUSCULUS CHROMOSOME 17 CLONE RP23-461H1, *** SEQUENCING IN
LOCUS PROGRESS ***; 12 unordered pieces.
DEFINITION AC151282
ACCESSION AC151282 GI:51511039
VERSION

KEYWORDS HTG; HTGS PHASE1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 188427)
WILSON,R.K.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188427)
AUTHORS Wilson,R.K.
TITLE Submitted
JOURNAL Direct Submission
Submitted (24-AUG-2004) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT ----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Project Information -----
Center project name: M_BA0461H01
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-primed ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 18266 bases at least Q40
Consensus quality: 183763 bases at least Q30
Consensus quality: 184341 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N' but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1129: contig of 1129 bp in length
1129: contig of 1129 bp in length
1230 1229: gap of unknown length
1230 2832: contig of 1603 bp in length
2832 2932: gap of unknown length
2932 4658: contig of 1726 bp in length
4658 4759: gap of unknown length
4759 6427: contig of 1669 bp in length
6427 6527: gap of unknown length
6527 11124: contig of 4597 bp in length
11124 11224: gap of unknown length
11224 21177: contig of 9953 bp in length
21177 21278: gap of unknown length
21278 31319: contig of 10042 bp in length
31319 31420: gap of unknown length
31420 54759: contig of 23339 bp in length
54759 54859: gap of unknown length
54859 95804: contig of 40946 bp in length
95804 95905: gap of unknown length
95905 135186: contig of 39282 bp in length
135186 135287: gap of unknown length
135287 135287: contig of 51397 bp in length
135287 186784: gap of unknown length
186784 186784: contig of 1644 bp in length.
Location/Qualifiers
1. 188427
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="17"
/clone="RP23-461H1"
1. 1129
/note="assembly_name:Contig11"
misc_feature
1230. .2832

misc_feature /note="assembly_name:Contig15"
2933. .4658
/note="assembly_name:Contig16"
misc_feature /note="assembly_name:Contig19"
4759. .6427
/note="assembly_name:Contig20"
6528. 11124
/note="assembly_name:Contig21"
11225. 21177
/note="assembly_name:Contig22"
21278. 31319
/note="assembly_name:Contig23"
31420. 54758
/note="assembly_name:Contig24"
54859. 95804
/note="assembly_name:Contig25"
95905. 135186
/note="assembly_name:Contig26"
135287. 186683
/note="assembly_name:Contig6"
186784. 188427
/note="assembly_name:Contig6"
ORIGIN
Query Match 72.0%; Score 21.6; DB 2; Length 188427;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 2 ATGCTGTTCACCGTTCGCTGTTC 29
Db 182832 ATGCTGTTCACCGTTCGCTGTTC 182859

RESULT 12
AC091123/c
LOCUS 129778 bp DNA linear PLN 09-AUG-2001
DEFINITION Oryza sativa chromosome 3 BAC OSJNB0093513 genomic sequence,
complete sequence.
ACCESSION AC091123
VERSION AC091123.4 GI:14670090
KEYWORDS HTG.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoae; Oryza.
1 (bases 1 to 129778)
Buell,C.R., Yuan,Q., Ouyang,S., Moffat,K.S., Hill,J.N.,
Gansberger,K., Brenner,M., Burgess,S., Hance,M., Shvartsbeyn,M.,
Teltrin,T., Riggs,E., Hsiao,J., Ziemann,V., Blunt,S., Pal,G.,
VanKen,S.E., Uteback,T.R., Feldlyum,T.V., Quackenbush,J.,
Salzberg,S.L., White,O. and Fraser,C.M.
Oryza sativa chromosome 3 BAC OSJNB0093513 genomic sequence
Unpublished
2 (bases 1 to 129778)
Buell,R.
REFERENCE Direct Submission
AUTHORS Submitted (31-MAR-2001) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
TITLE 3 (bases 1 to 129778)
Buell,R.
REFERENCE Direct Submission
AUTHORS Submitted (11-JUL-2001) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA
TITLE 4 (bases 1 to 129778)
Buell,R.
REFERENCE Direct Submission
AUTHORS Submitted (14-JUL-2001) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
TITLE 5 (bases 1 to 129778)
Buell,R.
REFERENCE Direct Submission
AUTHORS Submitted (18-JUL-2001) The Institute for Genomic Research, 9712
JOURNAL Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org

REFERENCE 6 (bases 1 to 129778)
AUTHORS Buell,R.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, rbuell@tigr.org
COMMENT On Jul 11, 2001 this sequence version replaced gi:1404299.
Address all correspondence to:rice@tigr.org

BAC clone OSJNB0093E13 is from Oryza sativa chromosome 3
The orientation of the sequence is from SP6 to T7 end of the BAC clone.
Genes were identified by a combination of several methods: Gene prediction programs including Fgenesh (<http://www.softberry.com/>), genscan and Genscan+ (Chris Burge, <http://CCR-081.mit.edu/GENSCAN.html>), GeneMarkHM (Mark Borodovsky, <http://genemark.biology.gatech.edu/GeneMark/>), and Genspicr (Mihela Pertea and Steven Salzberg, contact.miptea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (<http://www.tigr.org/tdb/tgi.html>). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as unknown proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). This BAC overlaps with rice BACs OSJNBa0091J19 (GB:AC084320) and OSJNB0024U04.

FEATURES

source

1. 129778
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="3"
/map="near S13014"
/clone="OSJNB0093E13"
/rpt_family="Gaijin Osz MITE element from gb:D10985 Rice 1701. .1821
repeat_region
complement(2452..2476)
repeat_region
/rpt_family="Gaijin Osz MITE element from gb:D10985 Rice 1701. .1821
repeat_region
complement(2452..2476)
repeat_region
/rpt_family="GAA)n"
4620..5435
/gene="OSJNB0093E13.4"
/note="predicted by fgenesh"
join(<4620..4718,4866..>5435)
/gene="OSJNB0093E13.4"
join(4620..4718,4866..5435)
/gene="OSJNB0093E13.4"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72893.1"
/db_xref="GI:14718315"
/translation="MEDWVLSDDSDSVLEHDGSDVSGSDSESGFAARSGKLPSSSQ LLLTLKCTQPKRTRRTLPVSPALAGSGHENAELSDVARNNDHVSNIAAEDTTF SGEDLDLNDIDECFDEBDGICENPDEIDFDEBEDDDSDSDIDDDIDECYDEVD ICEBNPDDEIDFDEBEIDCEBDLDDDDDCESLDNDNICFPAEDII CLETRTKMSST NIAM"
repeat_region
4640..4775
/rpt_family="(CGG)n"
repeat_region
complement(4730..4835)
/rpt_family="(CGG)n"
11008..11724
/gene="OSJNB0093E13.5"
/note="predicted by fgenesh"
<11008..>11724
mRNA
/gene="OSJNB0093E13.5"

CDS

11008..11724
/gene="OSJNB0093E13.5"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72895.1"
/db_xref="GI:14718317"

repeat_region

repeat_region
/rpt_family="(TGA)n"
complement(12766..16846)
/gene="OSJNB0093E13.7"

gene

GB:AMB82754 GI:2586082 (Oryza longistylis)atca)
complement(join(<12766..12903,13103..13518,13741..>16846))
/gene="OSJNB0093E13.7"
complement(join(12766..12903,13103..13518,13741..16846))
/codon_start=1
/product="putative gag-pol protein"
/protein_id="AAK72882.1"
/db_xref="GI:14718304"

CDS

translation="MASSSSSTAFANPLFGVQVSEKLTQNHSLSMAQVLTALRGARL EGHVLTGPPEALELEKEGKEKTVAVPNPAGWEPATQOVLGFSSLTREIRS OAVGPTAAAMKTIENPSTRSHAGATNVALTTQCGOSTVEYVSKRAAGDRI NATGKPTIDDEELVATITNGLDSEDPVVEALIAKASVTVAEVSQILGFENKVKIRI ACAATSGRSGGNGGNGRGRGTGGGGGGRGGGGRGNGRGGTDRNPTQVC HKKGIVNADCMHRVDENYDPBEKLGAAATNHYGVDTNNVOTEDATHITGDLKTRR EKYKGTQDIHTASGEGMDIOHIGHSYVTSRPLHKNILHVPKASKMLISVHRVAD NYAFLEIHQKFEILKDKYTRRTILEGPRRRLIPADDPILKOVAVNWSFRRMGR GASHKPYLVINONKILPCSNSESPSESCDCCOGKSNLPPEKFSVNDLEIHS DWGFPASDSVAKRIYSPFIDYKFWITITLKRKSRVFEKFEKQSVABEQFNKIL GMDTGWGEYQKLNSEFKQIGIAHQVSCPHHQOQGAVERHRHIVEIGLSLLAHASN PLKFWDEAFALATYILINRLPRKVIDPDLPLARLPHQDPYVSLRTPGCACPNLRPN THKLOFRSKQCVLFGFSNHLKGFCLAVATQVVISRPVTEDEGFPFSKIDSNVARI LRAEISLPLADLTSHASDGGVANSNDHNNILPVYTDSPENDADDTDQSAEAI NABOQDGIHPGASTGVHRAETNEVAASGPNRDRVATAPROAPATVEPAATSGAV SSRASRSIASHQIGDTPSPGPHHDAAGSSAAPEVIAATGCGDINARPKTRLOSGR KKKIYTDGTVRYGCTSSGEPONLAEALGDNNWEAMKEVYALMKNTMLHVPKRG INIIDCKWVYKIKRAADSLDRYKGRILVAKQFSGYSDYEDTSPPVKAATIRVLS LVASKDTSLPFYDNKGVFMILVYVDIDIVSSSEKAGYALDLDLQNGFALKDLDDLH PFLGIEYKKNVDGLVLEKAYSDIVLRVSGDCPVNTPISIKSKSOGGALLGPN DSTQGRSIVGALQVLTILRPDISFSVNNDPILVLCVHNKIKQPLDQVLTAKLTRHMQW QOKDCGYKHYFRNLV"
repeat_region
complement(13497..13708)
repeat_region
/rpt_family="Rice retrotransposon Tos7 DNA for reverse transcriptase, partial cds gi|1621482|dbj|D85871.1|D85871"
complement(16049..16146)
/rpt_family="(CGG)n"
17933..19954
/gene="OSJNB0093E13.6"
/note="predicted by fgenesh"
<17933..>19954
/gene="OSJNB0093E13.6"
17933..19954
/gene="OSJNB0093E13.6"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAK72884.1"
/db_xref="GI:14718306"
translation="MEDWVLSDDSDSVLEHDGSDVSGSDSESGFAARSGKLPSSSQ TPAIAYAVPQSPSPRPGFKYTSYGAFAFGIASSEHVAASSHAPVYDAABVRDAVS PVIYGEHENAEISDVASNDHVDNSINDAATVYTSRGEEDLDDETDGDLCEPDEED GICEBNPDDEIDFDEBEESDDEEDDGGSDLETSDSEYITESTDESDVDEBDTDLDES DSDBTBETESTSHDDELDDDDSDGSECPDEBDXITGENPDESDVDEBDTDLDES SDDDEDSYDSEIDDEBSDCDEIDDEBEHSGNKYDAIDNSFGDEESCMESQDA BEWMEPFGVAVSYNDIDTSDSMERIDGCKYDIDSESIYERBSVSDGSDDEBEERF AGGCTDGDIDYSLNGDDFREYLYQLADSGITNESFGDEESYLDDEVDLPHGINDLEFL DPFYGDITLYDYETESSCDECEHVCVCRKMWLIDGEEFYQLITGDEFDQULGSEIG GDAGADBEERPSDAGESDHDTPADGAGEAHGSDNMGASAAAAPASTSQQFOQ

	repeat_region	18749..18856	/rpe_family="(GGA)n"	
	repeat_region	complement(20115..20185)	/rpe_family="AT_rich"	
	gene	21515..21947	/gene="OSUNBb0093E13.8"	
		/note="predicted by tGeneSh"		
	mRNA	<21519..>21947	/gene="OSUNBb0093E13.8"	
	CDS	21519..21947	/gene="OSUNBb0093E13.8"	
		/codon_start=1		
		/product="hypothetical protein"		
Query Match	Similarity	70.7%; Score 21.2;	DB 8;	Length 129778;
Best Local	Matches	88.5%; Pred. NO. 2.2e+02;		
23;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps				0;
Db	20309	CATGTTAACCGTTCGTCTGTTCAA	20184	
RESULT_13				
AP006548	LOCUS	AP006548	140729 bp	DNA linear PLN 12-JUL-2003
DEFINITION	Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 9,			
SOURCE	PAC clone: P0705B1.			
VERSION	AP006548			
KEYWORDS	AP006548.1 GI:32490470			
ORGANISM	Oryza sativa (japonica cultivar-group)			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;			
	Echinozoidae; Oryzeae; Oryza.			
REFERENCE	1			
AUTHORS	Oliveira,A.C., Mattos,D.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,			
TITLE	Malone,G. and Dellagostin,O.			
JOURNAL	Oryza sativa nipponbare genomic DNA, chromosome 9, PAC			
AUTHORS	clone:P0705B1			
	Published Only in Database (2003)			
	2 (bases 1 to 140729)			
	Oliveria,A.C., Mattos,D.T., Carvalho,F.F., Shimano,A., Zimmer,P.D.,			
	Malone,G. and Dellagostin,O.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-JUL-2003) Antonio C Oliveira, Universidade Federal de			
	Pelotas, Centro de Genomica e Fitomelhioramento, Department			
	Fitotecnia; Campus Universitario do Capao do Leao, Pelotas,			
	Pelotas, Rio Grande do Sul 96010900, Brasil			
	(E-mail:acosole@terra.com.br,			
	url:http://www.nfpa.rche.br/faem/fitecna/fitomelhioramento/,			
	Tel:55-53-275-7263, Fax:55-53-275-9031)			
	Genes were predicted from the integrated results of the following:			
	GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), GENESEX			
	(http://www.softberry.com/), GeneMark.hmm			
	(http://opal.biology.gatech.edu/genemark/), Glimmer			
	(http://www.tigr.org/tdb/glimmer/glmr_form.html), RiceHMM			
	(http://isp.dna.afrc.go.jp/RiceHMM/), SplicePredictor			
	(http://bioinformatics.laestate.edu/cgi-bin/sp.cgi), BLASTN and			
	BLASTX. The genomic sequence was searched against NCBI Nonredundant			
	protein database, nr			
	(ftp://nbi.nlm.nih.gov/blast/db) and the EST sequence database at			
	RGP. Protein homologies of the coding regions were searched against			
	NCBI Nonredundant Protein database with BLASTP. ESTs represent the			
	identified cDNA sequences using BLASTN with the corresponding DDBJ			
	accession no. and RGP clone ID.			
	A gene with identity or significant homology to a protein is			
	classified based on the protein name to indicate the homology level			
	such as same name, 'putative-' and '-like protein'. A gene without			
	significant homology to any protein but with EST homology (covering			

almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to INGS standard.

The orientation of the sequence is from S86 to T7 of the PAC clone. This sequence of P0705E11 clone has an overlap with OSUNB0069216 clone (IDBJ: AP005862) at 5' end and an overlap with OSUNB0038K02 (DDBJ: AP005862) at 3' end.

location/Qualifiers
1..140728

source	location/Qualifiers
ORFS	1..140729 /organism="Oryza sativa (japonica cultivar-group)" /mol_type="genomic DNA" /cultivar="Nipponbare" /db_xref="taxon:39947" /chromosome="9" /clone="P0705E11" complement(1590..2675) /note="3'UTR"
gene	3382..4071 /gene="P0705E11.1" 3382..4071 /gene="P0705E11.1" /note="Probably inactive due to 3' exon missing in CDS probably inactive due to 5' exon missing in CDS pseudogene, orf3 of gypsy-type retrotransposon RIR2"
misc_feature	/pseudo complement(7148..9310) /gene="P0705E11.2" complement(7148..9310) /gene="P0705E11.2" /note="Probably inactive due to 5' exon missing in CDS pseudogene, Gag-Pol precursor of gypsy-type retrotransposon RIR2"
gene	/pseudo complement(13300..14385) /note="5'UTR"
LTR	20095..21408 /gene="P0705E11.3" <20095..21408 /gene="P0705E11.3" /note="start and end point are not identified"
gene	20095..21408 /gene="P0705E11.3" /codon_start=1 /product="putative anthranilate N-benzoyltransferase" /protein_id="BAC79154.1" /db_xref="GI:32490471"
ORFS	/translatability="MEVKVLSSKLVKPFVNGVAAADPEVYIPLSIFDKTYTKQMAI IYAPRPAPSTPAALIEKGLAAVLQAQYRAFAQOLSPDGEAAVNLNRGALVBAADA DLVMAAPAKTPPELLRLHPDLLEGLECVLLQLTRFKCSGLAVGFTSNHVAAGHATS NFLVAAAKATGTELPMGAPRVHNNHALFKRPPSPVHEHDHNRREYLPAADDSHGHD GGADIRIVYHKAHFTKQTFIAGLRAASEGRGPRSPRETTLAHMRTPMARGLSDE ASTRLSVDGGRHRLGAPAEYFGNVLVMAFPRAIVGDLITPLKHAQVHDEIVAPYD AYRPSFDLPALSGAGGKEGLAPSAVLKDVLCVPAEYDSWLTFFYELDGTSPYTF MPSYFPFEGMFLVPSYLDGGSVDAPVFNHNLIAFKKCYSMSE"
gene	33065..34423 /gene="P0705E11.4" <33065..34423 /gene="P0705E11.4" /note="start and end point are not identified"
mRNA	33065..34423 /gene="P0705E11.4" /codon_start=1 /product="putative anthranilate N-benzoyltransferase" /protein_id="BAC79155.1" /db_xref="GI:32490472"
CDS	/translatability="MEVKVLSRLVSPSPASAADEEFVPSMEDKTYTKQMAI IYAPRPAPSTPAALIEKGLAAVLQAQYRAFAQOLSPDGEAAVNLNRGALVBAADA DLVMAAPAKTPPELLRLHPDLLEGLECVLLQLTRFKCSGLAVGFTSNHVAAGHATS NFLVAAAKATGTELPMGAPRVHNNHALFKRPPSPVHEHDHNRREYLPAADDSHGHD GGADIRIVYHKAHFTKQTFIAGLRAASEGRGPRSPRETTLAHMRTPMARGLSDE ASTRLSVDGGRHRLGAPAEYFGNVLVMAFPRAIVGDLITPLKHAQVHDEIVAPYD AYRPSFDLPALSGAGGKEGLAPSAVLKDVLCVPAEYDSWLTFFYELDGTSPYTF MPSYFPFEGMFLVPSYLDGGSVDAPVFNHNLIAFKKCYSMSE"

[illegible]

```

gene
    gene      ILTNGKTPVEVLTPLQAAVNVOVKGLRPRTPKRNAHAKLSLLQCKMGOEFAERPDSE
              ILETFLORIAEEVEDDEHKRKKKEKITLGSLFALGRGH"
rnA          /gene="P0705B11.8"
             /join(=c5979.._66134,66228.._66279,67599.._75119,75210.._75
               75933.._76241,76395.._76475,76728.._77036,78187.._78573,
               78721.._79524,79613.._79975,80194.._80424,80691.._81035,
               81251.._81469,81516.._81572,81735.._82079,82295.._82513,
               82513)

Query Match       70.7%; Score 21.2; DB 8; Length 140729;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY     5 CTGTGCAACCGTTGGCTTGTCGA 30
        |||||                      |||||
Db     55011 CATGTTAACCGTGCCTTAATCCCA 55036

RESULT 14
LOCUS   AP005710                linear HTG 06-SEP-2002
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 9 clone P0646B07,
KEYWORDS  *** SEQUENCING IN PROGRESS ***.
ACCESSION AP005710
VERSION   AP005710.1 GI:22775481
KEYWORDS  HTG; HTGS PHASE2.
SOURCE    Oryza sativa (japonica cultivar-group)
           Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
           Spermatophytas; Magnoliophyta; Liliopsida; Poales; Poaceae;
           Ehrhartoidae; Oryzaceae; Oryza.
REFERENCE 1
AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE     Oryza sativa nipponbare(GAB3) genomic DNA, chromosome 9, PAC
JOURNAL   PUBLISHED ONLY in Database (2002)
REFERENCE 2 (bases 1 to 153675)
AUTHORS   Sasaki,T., Matsumoto,T. and Katayose,Y.
TITLE     Direct Submission
COMMENT    Submitted (05-SEP-2002) Takuji Sasaki, National Institute of
         Agrobiological Sciences, Rice Genome Research Program; Kamondai
         2-1-2, Tsukuba, Ibaraki,305-8602, Japan
         (E-mail:sasaki@nias.affrc.go.jp, URL:http://ryp.dna.affrc.go.jp/,
         Tel.:81-298-38-7441, Fax:81-298-38-7468)
NOTE     It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces believed them
to be correct as given, however the sizes of the gaps between them
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a "working draft" sequence.
* This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.

FEATURES
Source location/Qualifiers
1..153675
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="9"
/cloned=P0646B07"

ORIGIN
Query Match       70.7%; Score 21.2; DB 2; Length 153675;
Best Local Similarity 88.5%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY     5 CTTGTCAACC GTTCG TCTTGTCCA 30
        ||||||                  |||||
Db     19380 CATGTTAACCGTGCCTTAATCCCA 19405
```


RESULT 15
 LOCUS AP004118 155146 bp DNA linear PLN 21-JAN-2004
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 BAC clone:OJ1148_D05.
 ACCESSION AP004118
 VERSION AP004118.4 GI:41053257
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 2, BAC
 clone:OJ1148_D05
 Published Only in Database (2001)
 2 (bases 1 to 155146)
 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Direct Submission
 Submitted (29-APR-2001) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kamondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 E-mail:tsasaki@nias.affrc.go.jp, URL:http://rpg.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468
 On Jan 21, 2004 this sequence version replaced gi:32490504.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
 (http://www.softberry.com/), GlimmerM (http://opal.biology.gatech.edu/GlimmerM/), Glimmer
 (http://www.tigr.org/cdb/glimmer/glimmer.html), RiceHMM
 (http://rpg.dna.affrc.go.jp/RiceHMM/), SplicePredictor
 (http://bioinformatics.lasstate.edu/cgi-bin/sp.cgi), sim4
 (http://glabin.cse.psu.edu/html/docs/sim4.html), gap2
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI NonRedundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DBU. Protein homologues of the coding
 regions were searched against NCBI NonRedundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as
 a probable 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from -21m13 to M13rev of the BAC
 clone. This sequence of OJ1148_D05 clone has an overlap with
 OJ1743_B12 (DBJ: AP004096) clone at 5' end and with P0017H11
 (DBJ: AP004786) at 3' end. The sequence was generated by combining
 Monasato and RGP-Japan sequencing data. Detailed information on
 overlap and assembly quality together with annotation of this entry
 is available at
 http://rpg.dna.affrc.go.jp/GenomeSeq.html.
 Location/Qualifiers
 1. 155146
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1148_D05"
 5441..9517
 /gene="OJ1148_D05.1"

mRNA
 join(5441..5906,6423..6676,6786..6894,7393..7531,
 7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,
 9165..9517)
 /gene="OJ1148_D05.1"
 /note="supported by full-length cDNA(s) : AK105916"
 join(5463..5906,6423..6676,6786..6894,7393..7531,
 7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,
 9165..9517)
 /gene="OJ1148_D05.1"
 /note="supported by full-length cDNA(s) : AK101568"
 join(5577..5906,6423..6676,6786..6894,7393..7531,
 7646..7703,8300..8317,8422..8544,8646..8802,8884..8917,
 9165..9517)
 /gene="OJ1148_D05.1"
 /note="contains EST(s) : AU076051(E2768), AU076052(E2768)
 contains full-length cDNA(s) : AK105916, AK101568"
 /codon_start=1
 /product="putative silverleaf whitefly-induced protein 1"
 /protein_id="BAD07684.1"
 /db_xref="GI:41053258"
 /translation="MASPSLRDVAVGGLDRDPVSLIGKLGESRRLQNDPPALVPOED
 LVAQHVVDALPLVSMETGGPLVVRKVSVAEGSNIVYPTGPERAITSFVGMNDV
 VPANPEMPDFPSPFLPDSDDKLRGRTDCLGHVLAQIMRLRGVVKPKISV
 IAVFIANSNSLIITGIVDGLVADGLDKXGPLPMDPADKOPCIGTGIVTHLK
 AIKGLFHSGLAKKATINAMELMALKEIQTMTYNDPPRHEKELYPATPSITKPKTK
 STPGGLNDIPESCTISGDIRLPLPSTSVKQDEYVHDINENIEKLPTRKPSKY
 VLPDENLRGRLEITTDIEDIMNGVACNLESGRQALCKAEIEIVGHEVPYSITSLPI
 RELQDGFVQVTAQYGLIKTYAKNVCLEFSMAQGFVFLSIISQLEADV"
 9758..9946
 /gene="OJ1148_D05.2"
 <9758..9946
 /gene="OJ1148_D05.2"
 /note="start and end point are not identified"
 9758..9946
 /gene="OJ1148_D05.2"
 /note="predicted by GlimmerM etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD07685.1"
 /db_xref="GI:41053259"
 /translation="MRAQGAAREITSPSTFGQADGCAEGCGARARLPHFWPS
 TBARATARELTGRLPLYS"
 complement(10461..10670)
 /gene="OJ1148_D05.3"
 complement(<10461..>10670)
 /gene="OJ1148_D05.3"
 /note="start and end point are not identified"
 complement(10461..10670)
 /gene="OJ1148_D05.3"
 /note="predicted by GlimmerM etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD07686.1"
 /db_xref="GI:41053260"
 /translation="MPTIRPROPTLPQQLSLRGARVVSADBEERKPKDGKKRRGV
 GSRKRWLGDRNRGRGNEERTARR"
 11212..11718
 /gene="OJ1148_D05.4"
 <11212..>11718
 /gene="OJ1148_D05.4"
 /note="start and end point are not identified"
 11212..11718
 /gene="OJ1148_D05.4"
 /note="predicted by FGENESH etc."
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="BAD07687.1"
 /db_xref="GI:41053261"
 /translation="MTPRRRRPHDGLTRRSAGDDPAMPPTGSGRAVAARHSRLRV
 PLAVVELLATRTIGPTKIRSDSLDLPADGGRSPMGRGHPALPSRSSTPRARL
 SPCVYRHLMRREERKQGRQRMADDPFMALEGVYHFNSLGCLLEYLVGVSDFE
 EPIGGLIO"
 complement(join(15074..15178,17827..17877,17931..18029,

gene

gene

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:10:17 ; Search time 14.5242 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtgtcaacgctcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneeqn1980s:*
2: geneeqn1990s:*
3: geneeqn2000s:*
4: geneeqn2001s:*
5: geneeqn2001bs:*
6: geneeqn2002s:*
7: geneeqn2002bs:*
8: geneeqn2003s:*
9: geneeqn2003bs:*
10: geneeqn2003cs:*
11: geneeqn2003ds:*
12: geneeqn2004s:*
13: geneeqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	5 AAH76335	Aah76335 Z. mays M
2	30	100.0	255	5 AAH76340	Aah76340 Z. mays M
3	30	100.0	1394	2 AAH07408	Aax07408 Zea mays
4	30	100.0	1394	2 AAH07409	Aax07409 Zea mays
5	30	100.0	1394	5 AAH76332	Aah76332 Z. mays M
6	30	100.0	1394	5 AAH76333	Aah76333 Z. mays M
7	24.2	80.7	158	5 AAH76334	Aah76334 Z. mays M
8	22	73.3	1304	6 AB075351	Ab075351 Human lun
9	20.4	68.0	1542	6 AB213819	Ab213819 Arabidops
10	20.2	67.3	2357	8 ACA44013	Ac444013 Prokaryot
11	20	66.7	2000	12 ADJ40785	Adj40785 Plant CDN
12	19.6	65.3	2000	8 ADA71994	Ada71994 Rice gene
13	19.6	65.3	2000	8 ADA71994	Ada71994 Rice gene
14	19.6	65.3	2000	8 ADA71994	Ada71994 Rice gene
15	19.6	65.3	2000	8 ADA72530	Ada72530 Rice gene
16	19.6	65.3	2000	12 ADJ40977	Adj40977 Plant CDN
17	19.4	64.7	2240	2 AAQ47668	Aaq47668 Sequence
18	19.4	64.7	2243	4 AAF99956	Aaf99956 Rat brain
19	19.4	64.7	2243	9 ACD36214	Ac36214 cDNA enco
20	19.4	64.7	349980	6 ABQ81845	Abq81845 Bifidobac

C	21	19.2	64.0	2000	8 ADA73340	Ada73340 Rice gene
C	22	19	63.3	1324	8 ADA71777	Ada71777 Rice gene
C	23	19	63.3	1388	3 AAC44356	Aac44356 Arabidops
C	24	19	63.3	1821	10 ADE93879	Ade93879 Ixodes sc
C	25	19	63.3	5671	4 AAK90190	Aak90190 Human dir
C	26	19	63.3	5671	5 ABA20915	Ab20915 Human nlg
C	27	19	63.3	5671	19 AAS39832	Aas39832 Genomic s
C	28	19	63.3	5671	9 ADB32792	Ad32792 Human nov
C	29	19	63.3	198522	11 ACN44010	Acn44010 Human gen
C	30	18.8	62.7	883	6 ABK65288	Abk65288 Arabidops
C	31	18.8	62.7	883	10 ADD30938	Add30938 Plant yie
C	32	18.8	62.7	883	10 ADE37244	Ade37244 Plant yie
C	33	18.8	62.7	883	12 ADI41918	Adi41918 Plant tra
C	34	18.8	62.7	1083	4 AAS53644	Aas53644 Helicobac
C	35	18.8	62.7	1083	8 ACA34769	Ac34769 Prokaryot
C	36	18.8	62.7	1899	10 ABX06821	Abx06821 S. pneumo
C	37	18.8	62.7	1902	4 AAS55583	Aas55583 Streptoco
C	38	18.8	62.7	1902	8 ACA49918	Ac49918 Prokaryot
C	39	18.8	62.7	1908	12 ADJ63835	Adj63835 Plant 11p
C	40	18.8	62.7	2000	12 ADJ41535	Adj41535 Plant CDN
C	41	18.8	62.7	2247	13 ADS57809	Ads57809 Bacterial
C	42	18.8	62.7	2920	11 ADM02356	Adm02356 Human CDN
C	43	18.8	62.7	3339	10 ACF79841	Acf79841 Mouse ani
C	44	18.8	62.7	7812	3 AAZ61238	Aaz61238 DNA encod
C	45	18.8	62.7	9541	2 AAV52265	Aav52265 Streptoco

ALIGNMENTS

RESULT 1
AAH76335
ID AAH76335 standard; DNA, 30 BP.
AC AAH76335;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
PR (PION-) PIONEER HI-BRED INT INC.
PA
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
DR A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
PT
XX
XX Claim 5; Page 47; 50pp; English.
PS
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (II) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -152 to -181 bases upstream of

CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
 CC nucleotide sequence
 XX
 SQ Sequence 30 BP; 4 A; 9 C; 5 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.004;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30

RESULT 2

AAH76340
 ID AAH76340 standard; DNA; 255 BP.

AC AAH76340;

XX 29-OCT-2001 (first entry)

DE Z. mays Ms45 promoter fragment.

XX Ms45; male tissue; regulatory region; transcription; male fertility;

KM hybrid seed; promoter; de.

XX Zea mays.

OS WO200160997-A2.

XX 23-AUG-2001.

PD 13-FEB-2001; 2001WO-US004527.

XX 15-FEB-2000; 2000US-00504487.

XX (PION-) PIONEER HI-BRED INT INC.

XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;

XX WPI; 2001-514772/56.

XX A male tissue-preferred regulatory region comprising nucleotide sequences
 PT essential for initiating transcription of the Ms45 gene useful for
 PT mediating fertility in a male plant.

XX Example 5; Fig 8; 50pp; English.

XX The invention provides a male tissue-preferred regulatory region (I)
 CC comprising nucleotide sequences essential for initiating transcription of
 CC the Ms45 gene. A method of mediating male fertility in a plant is
 CC provided that involves introducing an expression vector comprising a
 CC promoter operably linked to (I) into a plant where the exogenous gene
 CC impacts male fertility of the plant and (I) controls expression of the
 CC exogenous gene. A method of producing hybrid seeds is also provided. The
 CC present sequence represents a Z. mays Ms45 promoter fragment
 XX
 SQ Sequence 255 BP; 59 A; 86 C; 39 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 5; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.0051;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

RESULT 3

AAK07408
 ID AAK07408 standard; DNA; 1394 BP.

AC AAK07408;

XX 08-JUN-1999 (first entry)

XX Zea mays Ms45 male tissue-preferred regulatory region.

DE Ms45; male; tissue-preferred; regulatory region; plant cells;

KM plant tissue; differentiated; maize; hybrid seed; fertility; ss.

XX Zea mays.

OS WO9859061-A1.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US012895.

XX 23-JUN-1997; 97US-00880499.

XX (PION-) PIONEER HI-BRED INT INC.

XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;

XX WPI; 1999-105628/09.

XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
 PT - useful in mediating plant fertility, especially hybrid seed production.

XX Claim 2; Page 22-23; 39pp; English.

XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
 CC region. It may be used in the construction of a vector for a method of
 CC producing exogenous genes in a male tissue- preferred manner, which is
 CC useful in restoring or conferring fertility, such as in hybrid seed
 CC production. In conferring fertility, a monocot/dicot plant is transformed
 CC with the exogenous nucleotide sequence (a male sterility gene, preferably
 CC Ms45), which encodes a product selected from auxins, rolb and dipteria
 CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
 CC and infertile plants
 XX

SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;

Query Match 100.0%; Score 30; DB 2; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0.0063;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
 |||||
 DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 4

AAK07409
 ID AAK07409 standard; DNA; 1394 BP.

XX AAK07409;

XX 08-JUN-1999 (first entry)

XX Zea mays Ms45 male tissue-preferred regulatory region.

DE Ms45; male; tissue-preferred; regulatory region; plant cells;

KM plant tissue; differentiated; hybrid seed; fertility; ss.

XX Zea mays.

OS WO9859061-A1.

XX 30-DEC-1998.

XX 19-JUN-1998; 98WO-US012895.

XX 23-JUN-1997; 97US-00880499.


```
XX (PION-) PIONEER HI-BRED INT INC.
PA Albertsen MC, Fox TW, Garnaat CW, Huffman GA, Kendall TL;
PI WPI; 1999-105628/09.
XX
XX New nucleic acid encoding a Ms45 male tissue-preferred regulatory region
PT - useful in mediating plant fertility, especially hybrid seed production.
XX
XX Claim 3; Page 23-24; 39pp; English.
XX
XX The sequence is that encoding an Ms45 male tissue-preferred regulatory
CC region. It may be used in the construction of a vector for a method of
CC producing exogenous genes in a male tissue- preferred manner, which is
CC useful in restoring or conferring fertility, such as in hybrid seed
CC production. In conferring fertility, a monocot/dicot plant is transformed
CC with the exogenous nucleotide sequence (a male sterility gene, preferably
CC Ms45), which encodes a product selected from auxins, rolb and dipheria
CC toxin. Hybrid seeds are produced by cross-pollinating maize male fertile
CC and infertile plants
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX ||||||||||||||||||||||||||||
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX RESULT 5
XX AAH76332 standard; DNA; 1394 BP.
XX
XX AAH76332;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
DE Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; de.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 4; Page 46; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 2; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX ||||||||||||||||||||||||||||
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX RESULT 7
```

```
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX ||||||||||||||||||||||||||||
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX RESULT 6
XX AAH76333 standard; DNA; 1394 BP.
XX
XX AAH76333;
XX
XX 29-OCT-2001 (first entry)
XX
XX Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
DE Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; de.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
PA Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
PI WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
XX Claim 4; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC tissue preferred regulatory region from Z. mays
XX
XX Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
SQ
XX
XX Query Match 100.0%; Score 30; DB 5; Length 1394;
XX Best Local Similarity 100.0%; Pred. No. 0.0063;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY
XX 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
XX ||||||||||||||||||||||||||||
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208
XX
XX RESULT 7
```



```

AAH76334
ID AAH76334 standard; DNA; 158 BP.
XX
AC AAH76334;
XX
DT 29-OCT-2001 (first entry)
XX
DE 2. may5 Ms45 male tissue-preferred regulatory region fragment.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
XX hybrid seed; db.
XX
OS Zea mays.
XX
PN MO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001MO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX WPI; 2001-514772/56.
XX
DR A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Claim 5; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
SQ Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
XX
Query Match 80.7%; Score 24.2; DB 5; Length 158;
Best Local Similarity 89.7%; Pred. No. 1.5;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTTCGTTGTTCC 29
DB 25 CATGCTGTTCACCGTTCGTTGTTCC 53

```

```

RESULT 8
ABQ75351/c
ID ABQ75351 standard; cDNA; 1304 BP.
XX
AC ABQ75351;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human lung specific nucleic acid sequence SEQ ID NO:90.
XX
KM Human; lung; lung specific nucleic acid; LSNA; lung specific protein;
XX LSP; cytostatic; gene therapy; vaccine; metastasis; lung cancer;
XX squamous cell carcinoma; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200264788-A2.

```

```

XX
PD 22-AUG-2002.
XX
PF 20-NOV-2001; 2001MO-US045080.
XX
PR 20-NOV-2000; 2000US-0252054P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Recipon H, Chen S, Sun Y, Liu C;
XX WPI; 2002-657601/70.
XX
DR New lung specific nucleic acid useful in gene therapy or as vaccines for
XX treating lung cancer (e.g. squamous cell carcinoma) or non-carcinoma lung
XX diseases, as well as for diagnosing, monitoring or staging these
XX diseases.
XX
PS Claim 1; Page 206; 282pp; English.
XX

```

```

CC The present invention describes an isolated lung specific nucleic acid
CC (LSNA) comprising a sequence that: (a) encodes any of the 93 amino acid
CC sequences comprising 17 - 733 amino acids, given in ABP52873 to ABP52965;
CC (b) comprises any of 115 sequences comprising 148 - 3193 base pairs (bp);
CC given in ABQ75262 to ABQ75376; (c) selectively hybridizes to (a) or (b);
CC or (d) has 60 % sequence identity to (a) or (b). LSNA and lung specific
CC protein (LSP) sequences have cytostatic activity and can be used in gene
CC therapy and vaccines. LSNA and LSPs are useful for diagnosing and
CC monitoring the presence and metastases of lung cancer in a patient. An
CC antibody that specifically binds to an LSP can be used for determining
CC the presence of an LSP in a sample, as well as for treating a patient
CC with lung cancer, particularly by inducing an immune response against the
CC lung cancer cell expressing the LSNA or LSPs. In particular, these LSNA
CC and LSPs are useful for identifying, diagnosing, monitoring, staging,
CC imaging and treating lung cancer (e.g. squamous cell carcinoma) and non-
CC carcinoma disease states in lung
XX
SQ Sequence 1304 BP; 352 A; 347 C; 314 G; 291 T; 0 U; 0 Other;
XX

```

```

Query Match 73.3%; Score 22; DB 6; Length 1304;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 CATGCTGTTCACCGTTCGTTGTTCCA 30
DB 1155 CTTTCTGTTCCTCCCGTTCCTTGTCCA 1126

```

```

RESULT 9
ABZ13819
ID ABZ13819 standard; DNA; 1542 BP.
XX
AC ABZ13819;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1624.
XX
KM Arabidopsis thaliana; plant; gene; stress; transgenic; db.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001MO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
XX
PR 26-JAN-2001; 2001US-0266467P.
XX
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI ) SCRIPPS RES INST.

```


PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 14; SEQ ID NO 1624; 577pp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1542 BP; 345 A; 382 C; 370 G; 445 T; 0 U; 0 Other;
XX
Query Match 68.0%; Score 20.4; DB 6; Length 1542;
Best Local Similarity 80.0%; Pred. No. 81;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
QY 1 CATGCTGTCAACCGTTCGCTTTGCCA 30
DB 237 CATGATGTTCACCAACATCGTCGTTCCA 266
XX
RESULT 10
ACA44013
ID ACA44013 standard; DNA; 2397 BP.
XX
AC ACA44013;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #25670.
XX
KM Antisense; ds; prokaryotic essential gene; cell proliferation;
KM drug design; gene.
XX
OS Pseudomonas putida.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU40143.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX
PS Claim 14; SEQ ID NO 31883; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2397 BP; 415 A; 825 C; 781 G; 376 T; 0 U; 0 Other;
XX
Query Match 67.3%; Score 20.2; DB 8; Length 2397;
Best Local Similarity 88.0%; Pred. No. 16+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1 CATGCTGTCAACCGTTCGCTTG 25
DB 516 CATGCTGTCAACCGTTCGCTTG 540
XX
RESULT 11
ADJ40785
ID ADJ40785 standard; cDNA; 2000 BP.
XX
AC ADJ40785;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant cDNA #1785.
XX
KM Plant; gene; ss; transcription; plant genome augmentation; cereal;
KM soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KM maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KM stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KM plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KM antifungal.
XX
OS Eukaryota.
XX
PN US2004016025-A1.
XX
PD 22-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00260238.
XX
PR 26-SEP-2001; 2001US-0325277P.
PR 26-SEP-2001; 2001US-0325448P.
XX


```

PR      04-APR-2002; 2002US-0370620P.
XX
XX      (BUDW/) BUDWORTH P.
PA      (MOUG/) MOUGHAMER T.
XX      (BRIG/) BRIGGS S P.
PA      (COOP/) COOPER B.
XX      (GLAZ/) GLAZEBROOK J.
PA      (GOFF/) GOFF S A.
XX      (KAT/) KATAGIRI F.
PA      (KREP/) KREPS J.
XX      (PROV/) PROVART N.
PA      (RICK/) RIQUE D.
XX      (ZHUT/) ZHU T.
XX
XX      Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI      Golf SA, Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;
XX
XX      WPI, 2004-190374/18.
XX
XX      New rice promoter, useful for manipulating crop plants to alter or
PT      improve phenotypic characteristics, e.g. produce large quantities of oil
PT      or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT      or high nutritional value.
XX
XX      Claim 26; SEQ ID NO 1785; 230bp; English.
XX
XX      The invention relates to plant nucleotide sequences that direct seed-,
CC      leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC      or constitutive transcription of an operatively linked nucleic acid
CC      segment. The invention also relates to a method for augmenting a plant
CC      genome and a method of identifying a gene, where its expression is
CC      altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC      in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
XX      canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
XX      sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC      encode are useful for manipulating crop plants to alter or improve
CC      phenotypic characteristics, to produce large quantities of oil or
CC      protein, to incur resistance to insecticides, viruses or fungi, and to
CC      incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC      have a high nutritional value with reduced apical dominance or dwarfism.
CC      early flowering or altered metabolic pathways. This sequence represents a
CC      plant nucleic acid of the invention. Note: The sequence data for this
CC      patent did not form part of the printed specification but was obtained in
CC      electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX      SQ      Sequence 2000 BP; 552 A; 438 C; 444 G; 566 T; 0 U; 0 Other;
XX
XX      Query Match      66.7%; Score 20; DB 12; Length 2000;
XX      Beet Local Similarity 82.1%; Pred. No. 1.2e+02;
XX      Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0
XX
XX      3 TCGTTGTTCAACCGTTGCGTCTGTTGCCA 30
XX      ||| ||||| ||||| ||||| |||||
DB      1179 TACATGTTTAACCGTTGCGTCTATTGCCA 1206
XX
XX      RESULT 12
XX      ID      ADA72061 standard; DNA; 2000 BP.
XX      AC      ADA72061;
XX      XX      ADA72061;
XX      DT      20-NOV-2003 (first entry)
XX      DE      Rice gene, SEQ ID 5386.
XX      KW      Plant; bacterial infection; fungal infection; viral infection; rice;
XX      gene; ds.
XX      XX      Oryza sativa.
XX      OS      Oryza sativa.
XX      PN      WO2003000898-A1.
XX

```

PD		03-JAN-2003.
XX		
PF	22-JUN-2001,	2001WO-IB001105.
XX		
PR	22-JUN-2001,	2001WO-IB001105.
XX		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
PB		
P1	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
XX	Katagiri F, Qian S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;	
DR	WPI; 2003-175290/17.	
XX		
PT	Identifying at least one gene involved in plant resistance or response to	
FT	pathogenic infection for conferring resistance or tolerance to a plant to	
PT	bacterial, fungal or viral infection by determining or detecting plant	
PT	gene expression.	
XX		
PS	Claim 27; SEQ ID NO 5386; 899pp; English.	
XX		
CC	The present invention relates to a method (M1) for identifying genes	
CC	involved in plant resistance or response to pathogenic infection. M1	
CC	comprises identifying a gene whose expression is significantly altered in	
CC	the incompatible interaction of plant gene expression relative to	
CC	expression of the gene in an uninfected plant, in a mutant plant that	
CC	does not express a gene associated with response to pathogenic infection,	
CC	or in a corresponding incompatible or compatible interaction. (M1) is	
CC	useful for conferring resistance to resistance or tolerance to a plant to	
CC	bacterial, fungal or viral infection. The present sequence was used to	
CC	illustrate the invention.	
XX		
SQ	Sequence 2000 BP; 605 A; 377 C; 326 G; 690 T; 0 U; 2 Other;	
	Query Match	65.3%; Score 19.6; DB 8; Length 2000;
	Best Local Similarity	84.6%; Pred. NO. 1.8e+02;
	Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0	
OY	5 CTGTTCACACCGTGCCTTTGTTCCA 30	
D8		
	974 CATGTTAACGGTGCTATTATCA 999	
RESULT 13		
ADA71994/C		
ID	ADA71994 standard; DNA; 2000 BP.	
XX		
AC	ADA71994;	
XX		
DT	20-NOV-2003 (first entry)	
XX		
DE	Rice gene, SEQ ID 5319.	
XX		
KW	Plant; bacterial infection; fungal infection; viral infection; rice;	
KW	gene; ds.	
XX		
OS	Oryza sativa.	
XX		
PN	WO2003000898-A1.	
XX		
PD	03-JAN-2003.	
XX		
PP	22-JUN-2001, 2001WO-IB001105.	
XX		
PR	22-JUN-2001, 2001WO-IB001105.	
XX		
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.	
XX		
P1	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
FI	Katagiri F, Qian S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;	
XX	WPI; 2003-175290/17.	
DR		
TX		

PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
XX gene expression.
PS Claim 27, SEQ ID NO 5319; 899bp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 638 A; 400 C; 347 G; 615 T; 0 U; 0 Other;
XX
Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 5 CTGTTCACCGTTCGCTCTGTTCGA 30
DB 150 CATGTTAACCGTTCGCTCTATTCA 125
XX
RESULT 14
ADA71736/c
ID ADA71736 standard; DNA; 2000 BP.
XX
AC ADA71736;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5061.
XX
KM Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
OS *Oryza sativa*.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PT 22-JUN-2001; 2001WO-IB001105.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI WPI; 2003-175290/17.
DR
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 27, SEQ ID NO 5061; 899bp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.

CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 601 A; 422 C; 306 G; 671 T; 0 U; 0 Other;
XX
Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 5 CTGTTCACCGTTCGCTCTGTTCGA 30
DB 291 CATGTTAACCGTTCGCTCTATTCA 266
XX
RESULT 15
ADA72530/c
ID ADA72530 standard; DNA; 2000 BP.
XX
AC ADA72530;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5855.
XX
KM Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
OS *Oryza sativa*.
XX
PN WO2003000898-A1.
XX
PD 03-JAN-2003.
XX
PT 22-JUN-2001; 2001WO-IB001105.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
PI WPI; 2003-175290/17.
DR
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
PS Claim 27, SEQ ID NO 5855; 899bp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 549 A; 413 C; 443 G; 588 T; 0 U; 7 Other;
XX
Query Match 65.3%; Score 19.6; DB 8; Length 2000;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
OY 5 CTGTTCACCGTTCGCTCTGTTCGA 30
DB 27 CATGTTAACCGTTCGCTCTATTCA 2

Search completed: September 15, 2005, 21:33:14
Job time : 18.5242 secs

Sat Sep 17 09:06:40 2005

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 14, 2005, 19:14:19 : Search time 4.53786 Seconds

(without alignments)
10817.505 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgtcaccgctcgtcgtcgtcga 30

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCBUS.COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	1394	US-08-880-499-1	Sequence 1, Appl1
2	30	100.0	1394	US-08-880-499-2	Sequence 2, Appl1
3	19.4	64.7	2243	US-07-937-609-15	Sequence 15, Appl1
4	19.4	64.7	2243	US-08-029-170-15	Sequence 15, Appl1
5	19.4	64.7	2243	US-09-443-745-15	Sequence 15, Appl1
6	19	63.3	45571	US-09-949-016-16262	Sequence 16262, A
7	18.8	62.7	7812	US-09-368-590-1	Sequence 1, Appl1
8	18.8	62.7	8756	US-09-949-016-1438	Sequence 1438, Ap
9	18.8	62.7	9541	US-08-961-527-132	Sequence 132, App
10	18.8	62.7	112874	US-09-949-016-13180	Sequence 13180, A
11	18.6	62.0	549	US-09-270-767-29480	Sequence 29480, A
12	18.6	62.0	11260	US-09-270-767-13492	Sequence 13492, A
13	18.4	61.3	1260	US-09-248-796A-3293	Sequence 3293, Ap
14	18.2	60.7	2710	US-09-573-080A-7	Sequence 7, Appl1
15	18	60.0	601	US-09-949-016-20052	Sequence 20052, A
16	18	60.0	601	US-09-949-016-20053	Sequence 20053, A
17	18	60.0	889	US-08-956-171E-1150	Sequence 1190, Ap
18	18	60.0	889	US-08-781-986A-1190	Sequence 1190, Ap
19	18	60.0	1591	US-08-728-956-3	Sequence 3, Appl1
20	18	60.0	111509	US-09-949-016-17379	Sequence 17379, A
21	18	60.0	254366	US-09-832-871-3	Sequence 3, Appl1
22	17.8	59.3	393	US-09-583-110-1388	Sequence 1388, Ap
23	17.8	59.3	393	US-08-107-433-743	Sequence 743, App
24	17.8	59.3	499	US-09-270-767-1993	Sequence 1993, Ap
25	17.8	59.3	499	US-09-270-767-17275	Sequence 17275, A
26	17.8	59.3	601	US-09-949-016-142937	Sequence 142937, A
27	17.8	59.3	601	US-09-949-016-204029	Sequence 204029, A

C 28	17.8	59.3	601	US-09-949-016-204030	Sequence 204030, A
C 29	17.8	59.3	978	US-08-858-207A-174	Sequence 174, App
C 30	17.8	59.3	999	US-08-961-527-193	Sequence 193, App
C 31	17.8	59.3	2009	US-07-958-222A-1	Sequence 1, Appl1
C 32	17.8	59.3	8073	US-09-949-016-12327	Sequence 12327, A
C 33	17.8	59.3	8074	US-09-949-016-16645	Sequence 16645, A
C 34	17.8	59.3	9472	US-08-325-547-9	Sequence 9, Appl1
C 35	17.8	59.3	40505	US-09-949-016-13439	Sequence 13439, A
C 36	17.8	59.3	72843	US-09-949-016-12574	Sequence 12574, A
C 37	17.8	59.3	141115	US-09-949-016-17490	Sequence 17490, A
C 38	17.8	59.3	183202	US-09-949-016-13614	Sequence 13614, A
C 39	17.8	59.3	265038	US-09-949-016-15779	Sequence 15779, A
C 40	17.8	59.3	276237	US-09-949-016-17504	Sequence 17504, A
C 41	17.4	58.0	331	US-09-640-211A-169	Sequence 169, App
C 42	17.4	58.0	601	US-09-949-016-178197	Sequence 178197, App
C 43	17.4	58.0	719	US-09-107-433-820	Sequence 820, App
C 44	17.4	58.0	789	US-09-252-991A-11034	Sequence 11034, A
C 45	17.4	58.0	813	US-09-252-991A-5455	Sequence 5455, Ap

ALIGNMENTS

RESULT 1
US-08-880-499-1Sequence 1, Application US/08880499
Patent No. 6037523

GENERAL INFORMATION:

APPLICANT: Albertson, Marc C.

APPLICANT: Fox, Tim W.

APPLICANT: Carl, Garnat W.

APPLICANT: Huffman, Gary A.

APPLICANT: Kendall, Timmy L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSER: PIONEER HI-BRED INTERNATIONAL, INC.

STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.

STREET: Box 1000

CITY: Johnston

STATE: Iowa

COUNTRY: USA

ZIP: 50131

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE: CONCURRENTLY HEREWITH

APPLICATION NUMBER: US/08/880,499

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Sweeney, Patricia A.

REGISTRATION NUMBER: 32,733

REFERENCE/DOCKET NUMBER: 0578

TELECOMMUNICATION INFORMATION:

TELEPHONE: (515) 248-4800

TELEFAX: (515) 248-4844

INFORMATION FOR SEQ. ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1394 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-880-499-1

Query Match 100.0%; Score 30; DB 3; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.0009;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/029,170
FILING DATE: 19930310
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-08-029-170-15

Query Match      64.7%; Score 19.4; DB 3; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTTGTCACCGCTGCTTGTTC 29
Db      5 CCGCTTGCTCACTGCTTGTTC 33

RESULT 5
US-09-443-745-15
Sequence 15, Application US/09443745
Patent No. 6706493
GENERAL INFORMATION:
APPLICANT: MANK, Stephen A.
TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,745
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
```

```

APPLICATION NUMBER: 08/029,170
FILING DATE: 10-MAR-1993
APPLICATION NUMBER: US 07/937,609
FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2243 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: rat brain CCKB receptor
FEATURE:
NAME/KEY: CDS
LOCATION: 136..1494
US-09-443-745-15
```

```

Query Match      64.7%; Score 19.4; DB 4; Length 2243;
Best Local Similarity 79.3%; Pred. No. 49;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      1 CATGCTTGTCACCGCTGCTTGTTC 29
Db      5 CCGCTTGCTCACTGCTTGTTC 33
```

```

RESULT 6
US-09-949-016-16262
Sequence 16262, Application US/09949016
Patent No. 6812338
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16262
LENGTH: 45571
TYPE: DNA
ORGANISM: Human
US-09-949-016-16262
```

```

Query Match      63.3%; Score 19; DB 4; Length 45571;
Best Local Similarity 81.5%; Pred. No. 13e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```


Qy	3	TGCTGTCAACCGTTCGTCCTCC	29
Db	38119	TGTTTGTTCATTTCGTTCTTCTTCC	38145

```

RESULT 7
US-09-368-590-1/C
? Sequence 1, Application US/09368590
? Patent No. 6187563
? GENERAL INFORMATION:
? APPLICANT: Solimena, Michele
? TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
? TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
? FILE REFERENCE: 101918-200 (OCR-941)
? CURRENT APPLICATION NUMBER: US/09/368, 590
? CURRENT FILING DATE: 1999-08-04
? EARLIER APPLICATION NUMBER: 60/095, 657
? EARLIER FILING DATE: 1998-08-07
? NUMBER OF SEQ ID NOS: 8
? SOFTWARE: FASTSEQ For Windows Version 3.0
? SEQ ID NO 1
? LENGTH: 7812
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)...(6879)
? NAME/KEY: unsure
? LOCATION: (100)...(102)
? NAME/KEY: unsure
? LOCATION: (1021)...(1023)
? NAME/KEY: unsure
? LOCATION: (2266)...(2268)
US-09-368-590-1

```

	Query Match	Similarity	Score	DB	Length
Best Local	812	76.7%	1.1e+02	DB 3	7812
Matches	23	Conservative	0	Mismatches	7
				Indels	0
				Gaps	0

```

RESULT 8
US-09-949-016-1438/C
: Sequence 1438, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1438
: LENGTH: 8756
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-1438

```

Query Match	62.7%	Score 18.8;	DB 4;	Length 8756;
Best Local Similarity	76.7%	Pred. No. 1.2e+02;		
Matches 23;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0

QY 1 CATGCTTGTTCAACCGTTGCTCTGTTCCA 30
||||| ||| ||| ||| |||
Db 1912 CATGCTTCTGCACACAGTGCTGTGCGCTTCA 1883

```

1      RESULT 9
2      US-08-961-527-132
3      ; Sequence 132, Application US/08961527
4      ; Patent No. 6420135
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Charles Kunsch
7      ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
8      ; NUMBER OF SEQUENCES: 391
9      ; CORRESPONDENCE ADDRESS:
10     ; ADDRESSEE: Human Genome Sciences, Inc.
11     ; STREET: 9410 Key West Avenue
12     ; CITY: Rockville
13     ; STATE: Maryland
14     ; COUNTRY: USA
15     ; ZIP: 20850
16     ; COMPUTER READABLE FORM:
17     ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
18     ; COMPUTER: HP Vectra 486/33
19     ; OPERATING SYSTEM: MSDOS version 6.2
20     ; SOFTWARE: ASCII Text
21     ; CURRENT APPLICATION DATA:
22     ; APPLICATION NUMBER: US/08/961,527
23     ; FILING DATE:
24     ; CLASSIFICATION: 424
25     ; PRIOR APPLICATION DATA:
26     ; APPLICATION NUMBER:
27     ; FILING DATE:
28     ; ATTORNEY/AGENT INFORMATION:
29     ; NAME: Brookes, A. Anders
30     ; REGISTRATION NUMBER: 36,733
31     ; REFERENCE/DOCKET NUMBER: PB340P1
32     ; TELECOMMUNICATION INFORMATION:
33     ; TELEPHONE: (301) 309-8504
34     ; TELEFAX: (301) 309-8512
35     ; INFORMATION FOR SEQ ID NO: 132:
36     ; SEQUENCE CHARACTERISTICS:
37     ; LENGTH: 9541 base pairs
38     ; TYPE: nucleic acid
39     ; STRANDEDNESS: double
40     ; TOPOLOGY: linear
41     ;
42     ; US-08-961-527-132

```

	Query Match	62.7%	Score 18.8	DB 3	Length 9541
	Similarity	76.7%	Pred. No. 1.2e+02		
b	Best Local	Matches 23	Conservative 0	Mismatches 7	Indels 0
	Gaps				0
Oy	1 CATGCTGTTCACCGTTCGTTGGTTCCA	30			
db	8049 CAGGCTTGTCCAACGTTCCATTTTTTCCA	8078			

```

RESULT 10
US-09-949-016-13180/C
; Sequence 13180, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01037
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```


Sat Sep 17 09:06:40 2005

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13180
; LENGTH: 112874
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(112874)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13180

Query Match          62.7%; Score 18.8; DB 4; Length 112874;
Best Local Similarity 76.7%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATGCTTGTCACCGCTGCTGTGTTCCA 30
DB 41167 CATGCTTGTCACAGGTGCTGCTCCA 41138

RESULT 11
US-09-270-767-29480/C
; Sequence 29480, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 29480
; LENGTH: 549
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-29480

Query Match          62.0%; Score 18.6; DB 4; Length 549;
Best Local Similarity 84.0%; Pred. No. 84;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGCTGCTTGT 27
DB 223 TGCTTGTGAGAGTCTTCTGT 199

RESULT 12
US-09-270-767-13492/C
; Sequence 13492, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 13492
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-13492

Query Match          62.0%; Score 18.6; DB 4; Length 1120;
Best Local Similarity 84.0%; Pred. No. 96;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGCTGCTTGT 27
DB 223 TGCTTGTGAGAGTCTTCTGT 199
```

```

RESULT 13
US-09-248-796A-3293
; Sequence 3293, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 3293
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-3293

Query Match          61.3%; Score 18.4; DB 4; Length 1260;
Best Local Similarity 78.6%; Pred. No. 1.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TGCTTGTTCACCGCTGCTTGTGTTCCA 30
DB 903 TGCTTGTTCACGTGTGTTCCGCCA 930

RESULT 14
US-09-573-080A-7/C
; Sequence 7, Application US/09573080A
; Patent No. 6828097
; GENERAL INFORMATION:
; APPLICANT: JOAN, KNOLL
; APPLICANT: ROGAN, PETER
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATT
; FILE REFERENCE: 30307
; CURRENT APPLICATION NUMBER: US/09/573,080A
; CURRENT FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 2710
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: repeat region
; LOCATION: (1)..(2710)
; OTHER INFORMATION: charlie3
; NAME/KEY: misc feature
; OTHER INFORMATION: n is a, c, g or t
; PUBLICATION INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Jurka, J; Walichiewicz, J; Milosavljevic, A
; TITLE: Prototypic sequences for human repetitive DNA
; JOURNAL: Journal of Molecular Evolution
; VOLUME: 35
; ISSUE: 4
; PAGES: 286-291
; DATE: 1992-10-
; DATABASE ACCESSION NUMBER: Database for repetitive elements (repbase)
; DATABASE ENTRY DATE: 1996-01-26
US-09-573-080A-7

Query Match          60.7%; Score 18.2; DB 4; Length 2710;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 20; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```


Qy 3 TGCTGTTCAACCGTTCGTTGTT 27
Db 1209 TGCTGTTTCACGAGTTCGTTGTT 1185

RESULT 15
US-09-949-016-200252/c
; Sequence 200252, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200252
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-200252

Query Match 60.0%; Score 18; DB 4; Length 601;
Best local similarity 80.8%; Pred. No. 1.6e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 5 CTGTTCAACCGTTCGTTGTTCCA 30
Db 239 CTGTTAAACAGTTCCTGTTGTTCCA 214

Search completed: September 15, 2005, 08:25:07
Job time : 7.53786 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11 ; Search time 19.5179 Seconds
(without alignment)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1179_1208

Perfect score: 30
Sequence: 1 catgctgttcaacgcttcgtctgttcca 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333128559 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	30	US-10-713-381-4	Sequence 9, Appli
2	30	100.0	255	US-10-713-381-9	Sequence 4, Appli
3	30	100.0	1394	US-10-713-381-1	Sequence 1, Appli
4	30	100.0	1394	US-10-713-381-2	Sequence 2, Appli
5	24.2	80.7	158	US-10-713-381-3	Sequence 3, Appli
6	22	73.3	1304	US-10-001-857-90	Sequence 90, Appli
7	21	70.0	1980090	US-10-719-993-6815	Sequence 6815, Ap

C	8	21	70.0	1980090	21	US-10-741-600-17676	Sequence 17676, A
	9	20.4	68.0	1542	9	US-09-938-842A-1624	Sequence 1624, Ap
	10	20.4	68.0	1542	11	US-09-938-842A-1624	Sequence 1624, Ap
	11	20.2	67.3	2397	17	US-10-282-122A-31883	Sequence 31883, A
	12	20	66.7	2000	17	US-10-260-238-1785	Sequence 1785, Ap
	C	13	19.6	65.3	640	US-10-437-963-96639	Sequence 96639, A
	C	14	19.6	65.3	901	US-10-425-115-3534	Sequence 3534, Ap
	C	15	19.6	65.3	2000	US-10-260-228-1977	Sequence 1977, Ap
	C	16	19.4	64.7	2243	US-09-443-745-15	Sequence 15, Appli
	17	19.4	64.7	2256646	19	US-10-470-565-1	Sequence 1, Appli
	C	18	19.5	63.3	1041	US-10-425-115-100318	Sequence 100318, A
	C	19	19.5	63.3	1821	US-10-425-563-5	Sequence 5, Appli
	C	20	19	63.3	1821	US-10-972-789A-5	Sequence 729, App
	C	21	19	63.3	5671	US-09-764-872-729	Sequence 729, App
	C	22	19	63.3	198522	US-10-087-192-244	Sequence 244, App
	C	23	18.8	62.7	883	US-09-934-455-279	Sequence 279, App
	C	24	18.8	62.7	883	US-10-225-068-183	Sequence 183, App
	C	25	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
	C	26	18.8	62.7	883	US-10-374-780A-381	Sequence 381, App
	C	27	18.8	62.7	883	US-10-225-068-183	Sequence 183, App
	C	28	18.8	62.7	883	US-10-225-068A-971	Sequence 971, App
	C	29	18.8	62.7	1083	US-09-815-242-7281	Sequence 7281, Ap
	C	30	18.8	62.7	1083	US-10-282-122A-22639	Sequence 22639, A
	C	31	18.8	62.7	1083	US-10-335-977-2194	Sequence 2194, Ap
	C	32	18.8	62.7	1125	US-10-335-977-2195	Sequence 2195, Ap
	C	33	18.8	62.7	1856	US-10-437-963-40048	Sequence 40048, A
	C	34	18.8	62.7	1859	US-10-472-928-2217	Sequence 2217, Ap
	C	35	18.8	62.7	1902	US-09-815-242-9220	Sequence 9220, Ap
	C	36	18.8	62.7	1902	US-10-282-122A-37788	Sequence 37788, A
	C	37	18.8	62.7	2000	US-10-260-238-2535	Sequence 2535, A
	C	38	18.8	62.7	2247	US-10-369-493-33483	Sequence 33483, A
	C	39	18.8	62.7	2220	US-10-108-260A-1041	Sequence 1041, Ap
	C	40	18.8	62.7	8788	US-10-756-149-64	Sequence 64, Appli
	C	41	18.8	62.7	9541	US-08-961-527-132	Sequence 132, App
	C	42	18.8	62.7	9541	US-10-158-844-132	Sequence 132, App
	C	43	18.6	62.0	2165598	US-10-472-928-4979	Sequence 4979, Ap
	C	44	18.6	62.0	375	US-09-974-300-3242	Sequence 3242, Ap
	C	45	18.4	61.3	174	US-09-814-353-5230	Sequence 5230, Ap

ALIGNMENTS

RESULT 1

US-10-713-381-4

Sequence 4, Application US/10713381

Publication No. US2004022331A1

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, MARC C.

APPLICANT: FOX, TIMOTHY W.

APPLICANT: GARNAT, CARL W.

APPLICANT: HUFFMAN, GARY

APPLICANT: KENDALL, TIMMY L.

TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF

TITLE OR INVENTION: USING SAME

FILE REFERENCE: 578R

CURRENT APPLICATION NUMBER: US/10/713,381

CURRENT FILING DATE: 2003-11-14

PRIOR APPLICATION NUMBER: 08/880,499

PRIOR FILING DATE: 1997-06-23

NUMBER OF SEQ ID NOS: 24

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 4

LENGTH: 30

TYPE: DNA

ORGANISM: Zea mays

US-10-713-381-4

Query Match 100.0%; Score 30; DB 20; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.0065;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTGTTCACGCTTCGCTGTTCACA 30

Db 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30

RESULT 2
US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 30; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB 39 CATGCTTGTTCAACCGTTCGTTGTTCCA 68

RESULT 3
US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 4

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match 100.0%; Score 30; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCCA 30
DB 1179 CATGCTTGTTCAACCGTTCGTTGTTCCA 1208

RESULT 5
US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 80.7%; Score 24.2; DB 20; Length 158;
Best Local Similarity 89.7%; Pred. No. 2.2;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CATGCTTGTTCAACCGTTCGTTGTTCC 29
DB 25 CATGCTTGTTCAACCGTTCGTTGTTCC 53

RESULT 6
US-10-001-857-90/c
; Sequence 90, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto


```

; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Pro
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/001,857
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 90
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-001-857-90
```

```
Query Match          73.3%; Score 22; DB 13; Length 1304;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
Db      1155 CTTCTTGTTCCTCCCGCTCTCTTGTTCCA 1126
```

```

RESULT 7
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 53342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815
```

```
Query Match          70.0%; Score 21; DB 20; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      2 ATGCTGTTCACCGTTCGCTTGTTCCA 30
Db      1948918 ATTCTTGCTGACAGTTCGCTTGTTCACCA 1948890
```

```

RESULT 8
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
```

```

; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676
```

```
Query Match          70.0%; Score 21; DB 21; Length 1980090;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      2 ATGCTGTTCACCGTTCGCTTGTTCCA 30
Db      1948918 ATTCTTGCTGACAGTTCGCTTGTTCACCA 1948890
```

```

RESULT 9
US-09-938-842A-1624
; Sequence 1624, Application US/0938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1624
; LENGTH: 1542
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1624
```

```
Query Match          68.0%; Score 20.4; DB 9; Length 1542;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY      1 CATGCTGTTCACCGTTCGCTTGTTCCA 30
Db      237 CATGATGTTCACCAATCGTCGTTTCCA 266
```

```

RESULT 10
US-09-938-842A-1624
; Sequence 1624, Application US/0938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
```


; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1624
 ; LENGTH: 1542
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-1624

Query Match	68.0%;	Score 20.4;	DB 11;	Length 1542;
Best Local Similarity	80.0%;	Pred. No. 1.2e+02;		
Matches 24; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;

```
QY      1 CATGCTGTTCACCGTTCGTCTTGTTCCA 30
          |||||
Db       237 CATGATGTTCAACCAATCGTCGGTTTCCA 266
```

RESULT 11
US-10-282-122A-31883
; Sequence 31883, Application US/10282122A
; Publication No. US20040029129A1

1. APPLICANT: Wang, Liangnu
2. APPLICANT: Zhang, Carlos
3. APPLICANT: Malone, Cheryl
4. APPLICANT: Haselbeck, Robert
5. APPLICANT: Ohlsen, Kari
6. APPLICANT: Zyckind, Judith
7. APPLICANT: Wall, Daniel
8. APPLICANT: Trawick, John
9. APPLICANT: Carr, Grant
10. APPLICANT: Yamamoto, Robert
11. APPLICANT: Forsythe, R.

```

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

```

;
;
; PRIOR APPLICATION NUMBER: 60/191,078
;
; PRIOR FILING DATE: 2000-03-21
;
; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR APPLICATION NUMBER: 60/230,347
 ;
 ; PRIOR FILING DATE: 2000-09-09
 ;
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ;

; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308

```

;
; Remaining Prior Application data removed
;
; NUMBER OF SEQ ID NOS: 78614
;
; SOFTWARE: PatentIn version 3.1

```

```

1  SEQ ID NO 31885
2  LENGTH: 2397
3  TYPE: DNA
4  ORGANISM: Pseudomonas putida

```

Query Match	Score
Best Local Similarity	67.3%
US-10-2882-122A-31883	88.0%: Pred.

Matches	22; conservative	0; M
QY	1 CATGCTGTTCAACCGTTCGTC	

Db 516 CATGCTGGTACAACCGTTCCGCTTG 540

RESULT 12
US-10-260-238-1785
; Sequence 1785, Application US/10260238
; Publication No. US20040016025A1

APPLICANT: Budworth, Paul R.
APPLICANT: Moughamer, Todd G.
APPLICANT: Briggs, Steven P.

; APPLICANT: Glazebrook, Jane
 ; APPLICANT: Goff, Stephen A.
 ; APPLICANT: Katagiri, Fumiyaki

```

;
; APPLICANT: Ricke, Darrell
;
; APPLICANT: Zhu, Tong
;
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
;
; FILE REFERENCE: 60111-NP
;

```

```

; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26

```

PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04

```

; SEQ ID NO 1785
; LENGTH: 2000
; TYPE: DNA

```

ORGANISM: *Oryza sativa*
US-10-260-238-1785

Query Match	66.7%	Score 20;	DB 17;	Length 2000;
Best Local Similarity	82.1%	Pred. No. 1.9e+02;		
Matches 23;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;

QY	3	TGCTTGTTCAACCGTTGCTTGTTCCA	30
Db	1179	TACATGTTAACCGTTGCTTTATTCAA	1206

RESULT 13
US-10-437-963-96639/c

Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.

APPLICANT:	NOVAKIC, DAVID
APPLICANT:	Zhou, Yihua
APPLICANT:	Cao, Yongwei
APPLICANT:	Wu, Wei

AFFILIANT: Bounadrov, Audrey A.
 ;
 APPLICANT: Barbazuk, Brad
 ;
 APPLICANT: Li, Ping
 ;
 TITLE OF INVENTION: Rice Nucleic

```

; FILE CONVENTION: FIDUCIS and USER INTERIOR FOR FIDUCIAL IMPROVEMENTS
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

```

```

: NUMBER OF SEQ ID NOS: 204966
: SEQ ID NO 96639
: LENGTH: 640
: TYPE: DNA

```

```

; ORGANISM: Oryza sativa
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94716C.1
US-10-437-863-86639

```

Query Match 65.3%; Score 19.6; DB 19; Length 640;

Best Local Similarity 84.6%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTGCTTCAACCGTTCGCTTGTCCA 30
Db 199 CATGTTTAAACCGTTCGCTTATTCAA 174

RESULT 14

US-10-425-115-3534/C
Sequence 3534, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yinhua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 3534
LENGTH: 901
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_103225C.1
US-10-425-115-3534

Query Match 65.3%; Score 19.6; DB 20; Length 901;
Best Local Similarity 84.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CATGCTTGTCAACCGTTCGCTTGT 26
Db 33 CACGCTTGTCAACCGTTCCTCTGT 8

RESULT 15

US-10-260-238-1977/C
Sequence 1977, Application US/10260238
Publication No. US20040016025A1
GENERAL INFORMATION:
APPLICANT: Budworth, Paul R.
APPLICANT: Moughamer, Todd G.
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Golf, Stephen A.
APPLICANT: Krepes, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Rieke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 6011-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/325,448
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/325,277
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 6077
SEQ ID NO 1977
LENGTH: 2000
TYPE: DNA
ORGANISM: Oryza sativa
US-10-260-238-1977

Query Match 65.3%; Score 19.6; DB 17; Length 2000;
Best Local Similarity 84.6%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 CTGCTTCAACCGTTCGCTTGTCCA 30
Db 291 CATGTTTAAACCGTTCGCTTATTCAA 266

Search completed: September 15, 2005, 20:45:45
Job time: 26.5379 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 08:25:31 ; Search time 134.222 Seconds
(without alignments)
14440.280 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatattccctcccccacacatccatccatccatgcaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_ptg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_dl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_str:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vt:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	40	100.0	40	6	AX224399 Sequence
2	40	100.0	50	6	AX224398 Sequence
3	40	100.0	158	6	AX224396 Sequence
4	40	100.0	255	6	AX224402 Sequence
5	40	100.0	1394	6	AX224394 Sequence
6	40	100.0	1394	6	AX224395 Sequence
7	40	100.0	1394	6	BD062176 Male t188
8	40	100.0	1394	6	BD062177 Male t188
9	40	100.0	1394	6	BD062178 Male t188
10	30.4	76.0	13737	8	AF360356 Zea may
11	24	60.0	55001	9	AC135206 Oryza sat
12	24	60.0	127196	10	AC084290 Homo sapi
13	24	60.0	171940	2	AL807784 Mouse Dna
14	23.6	59.0	193735	2	AC148514 Papio anu
15	23	57.5	117026	9	AC148501 Callithrix
16	23	57.5	149232	9	AC008897 Homo sapi
17	23	57.5	177104	2	AC102478 Mus muscu
18	23	57.5	195919	2	AC125876 Rattus no
19	23	57.5	207208	2	CR847932 Danio rer
					CR788256 Danio rer

20	23	57.5	218224	10	AC134404	AC134404 Mus muscu
21	23	57.5	234163	2	AC115318	AC115318 Rattus no
22	23	57.5	242109	2	AC147523	AC147523 Oryzomys
23	23	57.5	319056	2	AC016640	AC016640 Homo sapi
24	22.8	57.0	164944	3	AC008195	AC008195 Drosophila
25	22.8	57.0	175781	3	AC008091	AC008091 Drosophila
26	22.8	57.0	235928	3	AE003736	AE003736 Drosophila
27	22.6	56.5	95896	2	AL156384	AL156384 Human DNA
28	22.6	56.5	130893	2	AL149638	AL149638 Human DNA
29	22.6	56.5	160246	2	AC009221	AC009221 Homo sapi
30	22.6	56.5	161817	10	AC132119	AC132119 Mus muscu
31	22.6	56.5	170973	9	AP000923	AP000923 Homo sapi
32	22.6	56.5	175110	9	AP002428	AP002428 Homo sapi
33	22.6	56.5	186457	2	AC129669	AC129669 Rattus no
34	22.6	56.5	196990	2	AC134013	AC134013 Rattus no
35	22.6	56.5	235953	2	AC130117	AC130117 Rattus no
36	22.4	56.0	112361	9	H839817	AL023656 Human DNA
37	22.2	55.5	21700	9	BX322645	BX322645 Human DNA
38	22.2	55.5	71561	2	AL928982_6	Continuation (7 of
39	22.2	55.5	73885	2	AC025307	AC025307 Homo sapi
40	22.2	55.5	111122	8	ATFL1C1	AL132976 Arabidops
41	22.2	55.5	150489	2	AC069491	AC069491 Homo sapi
42	22.2	55.5	152797	2	AC016856	AC016856 Homo sapi
43	22.2	55.5	153814	2	AP003180	AP003180 Homo sapi
44	22.2	55.5	157574	9	AL592293	AL592293 Human DNA
45	22.2	55.5	176084	2	AC113354	AC113354 Homo sapi

ALIGNMENTS

RESULT 1	AX224399	Sequence 6 from Patent WO0160997.	40 bp	DNA	linear	PAT 10-SEP-2001
LOCUS	AX224399					
DEFINITION	AX224399					
ACCESSION	AX224399.1	GI:15554641				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
ORIGIN						
Query Match						
Best Local Similarity						
Matches						
QY						
Db						
RESULT 2						
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 5 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..50
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 4.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

Db 11 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 50

RESULT 3
AX224396 158 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 3 from Patent WO0160997.
AX224396
VERSION
AX224396.1 GI:15554638
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 3 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..158
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

Db 86 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 125

RESULT 4
AX224402 255 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 9 from Patent WO0160997.
AX224402
VERSION
AX224402.1 GI:15554644
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 9 23-AUG-2001;

AUTHORS
TITLE
JOURNAL

PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..255
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

Db 99 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 138

RESULT 5
AX224394 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0160997.
AX224394
VERSION
AX224394.1 GI:15554636
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 1 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40
|||||
1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

Db 1239 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 1278

RESULT 6
AX224395 1394 bp DNA linear PAT 10-SEP-2001
LOCUS
DEFINITION
Sequence 2 from Patent WO0160997.
AX224395
VERSION
AX224395.1 GI:15554637
KEYWORDS
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
1
Albertsen,M.C., Fox,T.W., Garnaat,C.W., Huffman,G. and Kendall,T.L.
Male tissue-preferred regulatory region and method of using same
Patent: WO 0160997-A 2 23-AUG-2001;
PIONEER HI-BRED INTERNATIONAL, INC. (US)

FEATURES
source
1..1394
/organism="Zea mays"
/mol_type="unassigned DNA"
/db_xref="taxon:4577"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 7
BD062176 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION
ACCESSION BD062176.1 GI:22607781
VERSION JP 2001520523-A/1.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1

REFERENCE
AUTHORS Albertsen, M.C., Fox, T.W., Garnaat, C.W., Huffman, G.A. and Kendall, T.L.
TITLE Male tissue-preferred regulatory region and method of using same
JOURNAL Patent: JP 2001520523-A 1 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/1

COMMENT

FEATURES
source 1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 8
BD062177 1394 bp DNA linear PAT 27-AUG-2002
LOCUS Male tissue-preferred regulatory region and method of using same.
DEFINITION
ACCESSION BD062177.1 GI:22607782
VERSION JP 2001520523-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1394)
Kendall, T.L.
Male tissue-preferred regulatory region and method of using same
Patent: JP 2001520523-A 2 30-OCT-2001;
PIONEER HI BRED INTERNATIONAL INC
PN JP 2001520523-A/2

COMMENT

PD 30-OCT-2001
PR 19-JUN-1998 JP 199504910
PR 23-JUN-1997 US 08/880499
PR MARC C ALBERTSEN, TIMOTHY W FOX, CARL W GARNAT, GARY A HUFFMAN,
PI TIMMY L KENDALL
PI TIMMY L KENDALL
PC C12N15/82, C12N15/29, C12N9/24, C12N9/22, C12N9/10, C12N9/00 PC
C07K14/34, C12Q1/68,
PC A01H5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
1. .1394
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 40; DB 6; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 40
|||||
Db 1239 AGGATACCTACTCCCAACAATCCATCTTACTCATGCAAC 1278

RESULT 9
AF360356 3343 bp DNA linear PIN 12-MAY-2001
LOCUS Zea mays male fertility protein (Ms45) gene, complete cds.
DEFINITION
ACCESSION AF360356
VERSION AF360356.1 GI:14028756
KEYWORDS Zea mays
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 3343)
Fox, T.W., Trimnell, M.R. and Albertsen, M.C.
Cloning of Ms45, a gene required for male fertility from Zea mays
Unpublished
2 (bases 1 to 3343)
Fox, T.W., Trimnell, M.R. and Albertsen, M.C.
Direct Submission
Submitted (13-MAR-2001) Trait and Technology Development, Pioneer
Hi-Bred Intl. Inc., 7300 N.W. 62nd Ave., P.O. Box 1004, Johnston,
IA 50131-1004, USA
Location/Qualifiers
1. .3343
/organism="Zea mays"
/mol_type="genomic DNA"
/culivar="B73"
/db_xref="taxon:4577"
/chromosome="9L"
<1392..>2942
/gene="Ms45"
join(<1392..1768,1898..2182,2280..2447,2534..>2942)
/gene="Ms45"
/product="male fertility protein"
join(1392..1768,1898..2182,2280..2447,2534..2942)
/gene="Ms45"
/codon_start=1
/product="male fertility protein"
/protein_id="AAK52489.1"
/db_xref="GI:14028757"
/translation="MEKNTLQWRGRDGIYQYPHLPFAALMALIYADPFGLSLAEV
DTRVYHILAPYGEVWGSPRDNASRLRGRLEFGEVGPESIFPDLOGRPYAGLA
DGRVVRMGEEAGMETFAVMNPDWSEVCANGVNSTTRKHQKEEFCGRPLGLRFGH
TEGLVADAYGLAVVQSGGVAASVAREADGDPTRFANDLDVHNGSVFFPDTSMR
SKDHLNTLIEGGTGRLRLYDPTSGVAVVLKGLVPRNGVQISDHQFLPSETTNC
RIMRYWLESPRAGEVVFANLPGPDPNVRNSRGQFWVALDCRTPADGVFAKRPWLR

TYLKEPLSLKVLTMKARMRHTVLLALDDEGRVVELEDRGHEVMKLVSEVREVGK
LWIGTAAHHNIATIPYPLED"

ORIGIN

Query Match 100.0%; Score 40; DB 8; Length 3343;
Best Local Similarity 100.0%; Pred. No. 2,6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAGCAATGCATCTTACTCATCGAAC 40
|||||
Db 1239 AGGATACCTACTCCCAAGCAATGCATCTTACTCATCGAAC 1278

RESULT 10

AC135206/c

AC135206 137327 bp DNA linear PLN 16-APR-2003
Oryza sativa (japonica cultivar-group) chromosome 3 clone
LOCUS OJ1041F02, complete sequence.

AC135206
AC135206.3 GI:27596977

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

/mol_type="genomic DNA"
/db_xref="caxon:33947"
/chromosome="3"
/clone="OJ1041F02"
/note="(japonica cultivar-group)"
complement(638..9495)
/gene="OJ1041F02.1"
/note="unknown protein"
complement(join(638..838,1141..1276,3132..3244,3326..3421,3513..3727,3812..3934,4024..4054,4157..4216,5088..5243,5322..5411,5484..5699,5808..5942,6035..6538,6622..6837,7548..7593,7679..7815,8196..8328,9221..9495))
/gene="OJ1041F02.1"
/codon_start=1
/product="unknown protein"
/protein_id="AAP06844.1"
/db_xref="GI:29893590"
/translation="MAQSLLELLIQFLMPNDARRQAEQIRRLARDPQVPALVHLLRTAKTPVROALAAVLLRKRTITSHWPKLPKAKSLKQALIDSTIDSHLVRAQANVVSIIKAVVAPGAEPELPIFOGSGPOEDREVALIFSLTETGTFQSHNDLQPIILKICLQDEASSRVRITALKVSGSFTEYNDGDVVKIPRDPSPSILNVRCLANGEEDVASTAFETPDELIESPAFLGDSVRSIVQSELEVCNDEINIRQDIIQITISM LVEKASFLKHKLVITPIQWCEPLTETABEDGSDLAADSAEVIDTMAINPRH VEPVLEPASVSFRHINKRYBAATSLGVVSEGCERLKKLEDCLVVLEALDOE QMVRGAAPFALQFAEHLQPEILSHYESVLICILNALRDPSEVKEKSYVALAACEE MGENILFYLDPLMCRVMSLOSPPNLOETGMSAIGVAAAQEMFVYAEVLEEMK GPMVLTNDEDCARARATEVGIAMAVGARMETILPPTIEAISGVLQVSELEXY THGFRSNVAEILDSFAYILPHVVPVLAASSCNDLGDGSAVIDDADSYNNGSGVSD D DVNDEPRVNRISMLTTIADILTAIRAITPAHADVLEKODILDTWNIYIKTWREDD KEVVAQACTSLADIVRDCFAIIEBYITRLADATILLRQSCCQGVSDGDDID HDDEVLMADVSLLPAFAKVMGSEYFDPITFLKFSIMKPKASBPPODTMVAATAEV AOGMGAPISAVYDKIMPLVLEKELASBATNRNAPCVGEMCKNGGAALAKYGDILH GSHRFADSEPDPAADNAGIARMIWOPQSIPLNQIILPVPRVIVNAPQVYVSPN ESDRYKTVAKAVASHLISVYGQNPILSLAPRANALASPRNR"
12372..15503
/gene="OJ1041F02.2"
/note="unknown protein"
join(12372..12757,12858..13025,13130..13295,13716..13826,13982..14130,14224..14354,14459..14558,15197..15503)
/gene="OJ1041F02.2"
/product="unknown protein"
/protein_id="AAP06845.1"
/db_xref="GI:29893591"
/translation="MPQGGRCRCHGMLGVGBAAVGGGEBFYVPLRKRLSVDGAS TAPRTICEGGEAGDITCDIYAAPLRSSCSAKAMPAPPLFRWMTTPPRRQGDGE EARRPGEAIRKGRKSYSLMLNQGISYSVGSALPQKLAASDPREKWTFRPP EGSKTPPHSHVDPRWKDYCPAVFRLHKLFGVDAEAYMLAICGDTLRELAASKPGSG CFPFTTQDDRPMIKTVKKSSEVKVILRMRSYVHVQYKSLTLTFRGTHCKKQAGCP KIHRRFDLKGSSGRTIDTERKIDETTLQDLQVAFRLQRFVYEELMKQIOMDCT FLETQGIWDYSLIGVHFRNDYSVKIGISQIHAPKSTGRKSKPEGSSPFCGLCFVE SCGCHRDILDSRKPRIQGINMPADAESSKIIIDNPLNBRHPIETPSSGSCCVYL FEGIIIDYDITKLEHAKYSFVNDYISAQDPKLYSRFPDPIRRVFIKEO"
16734..16878
/note="putative MITE, Galjin/Galjin-like"
19263..19454
/note="putative transposon, CACTA, En/Spm-like"
21819..24660
/gene="OJ1041F02.3"
/note="hypothetical protein"
join(21819..22007,22096..22206,23772..23996,24523..24660)
/gene="OJ1041F02.3"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAP06846.1"
/db_xref="GI:29893592"
/translation="WATASDSSGSGGEGORLLSPKEGERIAPTRRPDGTLRKAIRI RAGYPOBEVALIYSGKQMRKSGPGVDPALDAKPTKAAKRNRRKRRKQOAS TTNDGKLHIEDGERTNDKADVSTKQISGAIASISLVVATSTDAFNSSES SAPDIDKIRALKKIRLAENAVQDPENLAKPEQLEKKKIBGMKEBLKLENKSPA AS"
complement(26563..30575)
/gene="OJ1041F02.4"

FEATURES
Source
1..137327
/organism="Oryza sativa (japonica cultivar-group)"

CDS
 /note="unknown protein"
 complement(join(26563..26772,27082..27288,27430..27537,
 27881..27940,28051..28126,28327..28484,28616..28678,
 28790..28874,29560..30575))
 /gene="OJ1041F02.4"
 /codon_start=1
 /product="unknown protein"
 /protein_id="AAP06847.1"
 /db_xref="GI:29893593"
 /translation="MGPPAAPSPSPSSSGSSRRRLRERNAKHLIGDASNFCAY
 QSPPAASAPASGSPSLACSPACSLDLTFPIGSGDCDQVQLCSLIGSDVDAV
 PADMEAHKAKRSFSSSSASTPKRPEPPADSPVREVAEPEPLPPAPAPLVP
 AKTPREVAIEAPILRVDPWEPPAPDVAKSGEGIKVAREPPPEPPWAPA
 VCVESTWDLRSAPREDESHAHAPASRSGDSCADQAGEEDDAAVLTLELRICE
 TSEEFSTSLSTNDDEISITTESMFYISPMGRFPRKIRSNRMGLGSGSGFVPE
 EGISEGVFPFAVKEVCLCDGSAQOCIQLQEOLALSQPEHENVQYGTDKESK
 LYFLFLVQSGSLASLYQKRYLRDTHASATROIINGLTALHERNIYHRIKANTLV
 HANGSVLADFGLAKEITKFNVLKSCGYVMAPEVNPVKTTYGPEADTWSLGCYVL
 EMLTROLPYGLEMTOLVRIKGPEVILICSPRADRISOCYKPNQDRESAKL
 LEHPVNRMSRISRMKTSRNSISLVLIMMSPEORTIRREALISTKQSTPTTIA
 FHLILSHSLRLGDMWFSQRLIFSPENRLNVHSSA"
 repeat_region
 /note="putative MITE, MITE-adj, type D-like"
 complement(31668..31815)
 repeat_region
 /note="putative MITE, MITE-adj, type D-like"
 complement(35049..49855)
 gene
 /gene="OJ1041F02.5"
 /note="hypothetical protein"
 complement(join(35049..35184,35866..36033,36922..37205,
 37280..38523,40038..40104,40196..40287,40406..40617,
 41678..41760,42229..42259,42907..43088,43694..43613,
 44329..44426,44622..44676,45877..45992,46149..46313,
 46447..46620,47271..47470,47915..48019,48200..48263,
 49589..49855))
 /gene="OJ1041F02.5"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAP06848.1"
 /db_xref="GI:29893594"
 /translation="MGGSADPEAPPTPTSPSPAPAKATPSPASADGNRLRVCQSLKS
 MGPPKAGGGGAGAGLPLAAGDGPPEVKRGRPRKSEAGKESNRETTGLDQD
 SKDEVILVDESPOKORKGRKQKGAALKVPRKHCALKESTGHSQCOOLRESOTD
 VLPOKSTVDIDLVTGPSASPVNDVADLNDNEKPOLIVLDRSANTLAENRRLS
 SKCKMHPASRKIKGACODILANVEDEDSCAREPDRPCPVYVYELRYTMPIH
 WSKMLADSKSLGTSTTEBNSAEHADPEGHNLNFKDKKSPSSQDVIVDDECL
 ASSSCFHASLFESKQHERVQHELPEVTPKCCOTANIMTKRYETALNSHSPSICY
 LNMGLGMDGERHNRKONITVNGSNVGRSCODSDTDYSADADENVLITPVCQ
 KSAAPFACAEQGFNYELNMPKTVVPPPLPHNKLAAOTLDLDPDKLHNSHSICYI
 PTRFVHMNRNIIISIAHSSSLSVIKHCLVNTSMRNGAVYRQFEETKSGHLEK
 WQGEETILGPSISLSDPAGCTPGTAETKQYINKTILFEDVDPVPEDEGRFISTIKM
 VETTKPPIITLSNKDPPPLHLAQLVLDITPSSABLSHVMIKCSGEVETIVQ
 KHLIDAFGLRLKELCPSLLDADAVSTPRIMPDPCKSETIYMEIDKTIIVAE
 OKKQWESVEFEGLEIQIMTLPKGRSAGTRPKSKLKGHSADCNDSAPKNDL
 DPHSDPDIPLPSNORMNRNRGVVLPFAESDDLDADAAADATFTVOEGRLPQSEEL
 PCLYGHSINIVSEVFPQSSVPLHREVTSNOLCFPSPEPAPFEFASPNONLESNM
 RGSISQICPFMSOGTSCVPSFPMVCGTSASISDDLSLIVNSGLAHRNSETTAA
 SVASLEDTKVENQMTDKQCKMEDEYGERICEYVELADNHRSSGSLITGYLMDSCS
 RAESVWLISGKNNDSCKEHVQDVTNMRKQCHPVLCDMNNHRSVGLAKRYSRQ
 LISEDMLMLISCHPFSNDISDPLITPTTESDGSYSKOLEMGSIVYOHGICLFIQSOQ
 ATSDGDFLQELIFSGTTTSLGKTFVSGISCGDSGNSHVKYTCISCRKREDA
 RLAEVLVPPVPRKLSLRGPAFVDVLYSSMSOISOLENNOLSCKASSQRRRCQRH
 YLSSGALSIASRIGLACSTFSDRESSTITEQALS"
 repeat_region
 /note="putative MITE, MITE-adj, type M-like"
 complement(39012..39207)
 repeat_region
 /note="putative MITE, MITE-adj, type M-like"
 complement(39681..39879)
 repeat_region
 /note="putative MITE, MITE-adj, type G-like"
 complement(51474..51836)
 gene
 /note="putative MITE, Castaway-like"
 complement(59793..62196)
 /gene="OJ1041F02.6"
 Query Match 76.0%; Score 30.4; DB 8; Length 137327;
 Best Local Similarity 85.0%; Pred. No. 0.12;
 Matches 34; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AGAATACCTACTCTCCCAACATCATCTTACTCATCAAC 40
 Db 119478 AGAACACCTACTCCCAACATCATCGCTGATCATCAAC 119439
 RESULT 11
 AC084290/c 55001 bp DNA linear PRI 28-MAR-2002
 LOCUS AC084290/c
 DEFINITION Homo sapiens 12 BAC RP11-568G5 (Rosewell Park Cancer Institute Human
 BAC library) complete sequence.
 AC084290
 VERSION AC084290.15 GI:19774277
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 55001)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banke,T.,
 Barbarella,J., Benton,J., Bimaye,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhey,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chiu,D., Chowhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,
 Escotto,M., Falls,T., Ferraguto,D., Flaegs,N., Ford,J., Foster,P.,
 Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N.,
 Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Hawlik,P.,
 Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M.,
 Hollaway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Huliyk,S.,
 Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,
 Johnson,K.R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U.,
 King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N.,
 Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O.,
 Lieu,C., Liu,J., Liu,W., Loulsegod,H., Lozado,R.J., Lu,X.,
 Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapun,P.,
 Mardiney,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,
 Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Mescher,S.,
 Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,
 Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,
 Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwunu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojiboken,I., Roife,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shm,C., Shoshari,N., Slaton,I.,
 Sodergren,E., Soneike,T., Sparks,A., Stanley,H., Stone,H.,
 Suton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tamey,J., Taylor,C., Taylor,T., Telitod,B., Thomas,N., Thomas,S.,
 Umanui,K., Vaequez,L., Vera,V., Villalona,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlingon,S.,
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kuchelapati,R.,
 Weinstein,G., and Gibbs,R.
 Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 55001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (21-OCT-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 55001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (27-MAR-2002) Human Genome Sequencing Center, Department

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 55001)

Worley, K.C.

Direct Submission

Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On Mar 28, 2002 this sequence version replaced gi:19747063.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

FEATURES
Source Location/Qualifiers

1. .55001
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-568G5"
1. .1301
/note="overlaps bases 177088..178388 of clone AC079630"
/function="clone overlap"
35. .229
repeat_region
/rpt_family="MIR"
1002. .1023
repeat_region
/rpt_family="AT_rich"
1092. .1227
repeat_region
/rpt_family="MIR"
complement(2156..2209)
/rpt_family="L2"
repeat_region
3563. .3650
repeat_region
/rpt_family="(CCGTA)n"
complement(3688..3765)
repeat_region
/rpt_family="L2"
complement(3840..3908)
repeat_region
/rpt_family="L1MD3"
complement(4234..4464)
repeat_region
/rpt_family="MIR"

repeat_region complement(4847..4914)
repeat_region /rpt_family="MIR"
4938..5053
repeat_region /rpt_family="L2"
6330..6660
repeat_region /rpt_family="AluSg"
complement(7309..7743)
repeat_region /rpt_family="LTR43"
7744..9144
repeat_region /rpt_family="L1PA4"
complement(9151..9317)
repeat_region /rpt_family="LTR43"
10587..10846
repeat_region /rpt_family="L1M4"
10912..11015
repeat_region /rpt_family="L1P"
11096..11271
repeat_region /rpt_family="L1ME"
11450..11635
repeat_region /rpt_family="MLT11"
11941..12179
repeat_region /rpt_family="MIR"
12797..12828
repeat_region /rpt_family="(CATATA)n"
complement(12975..13087)
repeat_region /rpt_family="FLAM_C"
13714..13737
repeat_region /rpt_family="AT_rich"
14058..14219
repeat_region /rpt_family="MER45"
14411..14545
repeat_region /rpt_family="MIR"
complement(15058..15127)
repeat_region /rpt_family="MIR"
16186..16413
repeat_region /rpt_family="MIR"
complement(17624..17922)
repeat_region /rpt_family="AluUb"
complement(18309..18606)
repeat_region /rpt_family="AluSx"
18644..18854
repeat_region /rpt_family="AT_rich"
complement(18855..19135)
repeat_region /rpt_family="AluUb"
19136..19158
repeat_region /rpt_family="AT_rich"
complement(20660..20753)
repeat_region /rpt_family="MER91C"
22179..22680
repeat_region /rpt_family="(TAAA)n"
22180..22358
repeat_region /rpt_family="AluJo"
22355..22632
repeat_region /rpt_family="AluUb"
22633..22680
repeat_region /rpt_family="(TAAA)n"
22682..22789
repeat_region /rpt_family="AluJo"
22796..23099
repeat_region /rpt_family="AluY"
complement(23147..23327)
repeat_region /rpt_family="MIR"

Query Match 60.0%; Score 24; DB 9; Length 55001;
Best Local Similarity 84.4%; Pred No. 48;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATTCATCTTACTCAT 35
DB 47598 AAACCTTACTCCCAACATTCATCTTACTCAT 47567

RESULT 12

AL807784/c
LOCUS AL807784 127196 bp DNA linear ROD 13-NOV-2002
DEFINITION Mouse DNA sequence from clone RP23-448C18 on chromosome X, complete
sequence.
ACCESSION AL807784
VERSION AL807784.11 GI:25045332
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS 1 (bases 1 to 127196)
TITLE Howden, P.
JOURNAL Direct Submission
Submitted (13-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone request: clonerequest@sanger.ac.uk
On Nov 15, 2002 this sequence version replaced gi:24939941.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-448C18 is
from the RPCI-23 Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
Location/Qualifiers
FEATURES
source
1. 127196
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP23-448C18"
/clone_1b="RPCI-23"

ORIGIN
Query Match 60.0%; Score 24; DB 10; Length 127196;
Best Local Similarity 75.0%; Pred. No. 43;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
1 AGGATACCTACTCCCAACATCATCTACTCATGCAAC 40
Db 110413 AGGATACATAGCCAAATCACTTCTTACTAATGCCAC 110374

RESULT 13
AC148514/c
LOCUS AC148514 171940 bp DNA linear HTG 24-MAR-2004

DEFINITION Papio anubis clone RP41-375M17, WORKING DRAFT SEQUENCE, 3 ordered
pieces.
AC148514
AC148514.2 GI:45680463
VERSION HTG; HTGS PHASE2; HTGS DRAFT.
KEYWORDS Papio anubis (olive baboon)
SOURCE Papio anubis
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
REFERENCE
AUTHORS 1 (bases 1 to 171940)
TITLE Antionellis, A., Ayele, K., Benjamin, B., Blakestey, R.W.,
Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
Coleman, H., Daki, N., Engle, J., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B.,
Idol, J.R., Jones, C., Karlins, E., Kim, H., Kwong, P., Latic, P.,
Larsen, S., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masello, C., Maskeli, B., McDowell, J.,
Mulikin, J.C., Paguirigan, C., Portnoy, M.E., Prasad, A., Puti, O.,
Reddix-Dugue, N., Schandier, K., Schueler, W.G., Shah, K., Sison, C.,
Stantipod, S., Thomas, J.W., Thomas, P.J., Tsipouri, V., Vogt, J.L.,
Wetherby, K.D., Young, A. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 171940)
Green, E.D.
Direct Submission
Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
Grovemont Circle, GaitHERsbury, MD 20877, USA
3 (bases 1 to 171940)
Green, E.D.
Direct Submission
Submitted (24-MAR-2004) NIH Intramural Sequencing Center, 8717
GroveMont Circle, GaitHERsbury, MD 20877, USA
On Mar 24, 2004 this sequence version replaced gi:45332443.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nih.gov
----- Project Information
Center project name: f5m
Center clone name: 375M17

The sequence data in this record represents an 'enhanced'
version of a phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8x average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 171395 bases at least Q40
Consensus quality: 171621 bases at least Q30
Consensus quality: 171701 bases at least Q20
Insert size: 173000; agarose-fp
Insert size: 171740; sum-of-configs
Quality coverage: 8.27x in Q20 bases; agarose-fp
Quality coverage: 8.33x in Q20 bases; sum-of-configs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have

* provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 10148: contig of 10148 bp in length
 * 10149 10248: gap of unknown length
 * 10249 103023: contig of 92775 bp in length
 * 103024 103123: gap of unknown length
 * 103124 171940: contig of 68817 bp in length.
 Location/Qualifiers

FEATURES

source

1..171940
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /db_xref="taxon:9555"
 /clone="RP41-375M17"
 /clone_1lb="RP41"
 /note="BAC resource: http://bacpac.chori.org/"
 1..10148
 /note="assembly_fragment"
 clone_end:SP6
 vector_side:left"
 10249..103023
 /note="assembly_fragment"
 103124..171940
 /note="assembly_fragment"
 clone_end:T7
 vector_side:right"

ORIGIN

Query Match 60.0%; Score 24; DB 2; Length 171940;
 Best Local Similarity 84.4%; Pred. No. 42;
 Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCCATCTTACTCAT 35

DB 33556 AAACCTTACTCCCAACATCTTACTCAT 33525

RESULT 14

AC148501/c

LOCUS Callitrix jacchus clone CH259-368E20, WORKING DRAFT SEQUENCE, 11
 DEFINITION Ordered pieces.

ACCESSION

AC148501

AC148501.3 GI:50080339

HTG; HTGS PHASE2; HTGS DRAFT.

KEYWORDS

SOURCE

ORGANISM

Callitrix jacchus (white-tufted-ear marmoset)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
 Callitrix.

REFERENCE

AUTHORS

1 (bases 1 to 193735)
 Antorellis A., Aylee, K., Benjamin, B., Blakeley, R.W.,
 Bouffard, G.G., Brinkley, C., Brooks, S., Chu, G., Coleman, B.,
 Coleman, H., Daki, N., Engle, D., Guan, X., Gupta, J., Haghighi, P.,
 Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Jones, C.,
 Karlins, E., Kim, H., Kwong, P., Latic, P., Larson, S., Lee-Lin, S.-O.,
 Legaspi, R., Madden, M., Maduro, O.L., Maduro, V.B., Margulies, E.H.,
 Masello, C., Maskeri, B., McDowell, J., Mullikin, J.C., Peguirigan, C.,
 Park, M., Portnoy, M.E., Prasad, A., Puri, O., Reddix-Dugue, N.,
 Schandler, K., Schueler, M.G., Shah, K., Sison, C., Stantirip, S.,
 Thomas, D.W., Thomas, P.J., Tsipouri, V., Vogt, J.L., Wetherby, K.D.,
 Young, A. and Green, E.D.
 NISC Comparative Sequencing Initiative

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (10-MAR-2004) NIH Intramural Sequencing Center, 8717
 Grove Mount Circle, Gaithersburg, MD 20877, USA

REFERENCE

AUTHORS

Green, E.D.
 Direct Submission

JOURNAL

COMMENT

Submitted (09-JUL-2004) NIH Intramural Sequencing Center, 8717
 Grove Mount Circle, Gaithersburg, MD 20877, USA
 On Jul 9, 2004 this sequence version replaced gi:46430810.

----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@nigr.nih.gov
 ----- Project Information
 Center project name: ffw
 Center clone name: 368E20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones; overlaps with neighboring clones; alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.99019
 Consensus quality: 190717 bases at least Q40
 Consensus quality: 191709 bases at least Q30
 Consensus quality: 193323 bases at least Q20
 Insert size: 14400; agarose-fp
 Insert size: 192735; sum-of-contigs

Quality coverage: 13.33x in Q20 bases; agarose-fp
 Quality coverage: 9.96x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 11 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 29309: contig of 29309 bp in length
 * 29310 29409: gap of unknown length
 * 29410 37712: contig of 8303 bp in length
 * 37713 37812: gap of unknown length
 * 37813 40472: contig of 2660 bp in length
 * 40473 40572: gap of unknown length
 * 40573 45724: contig of 5152 bp in length
 * 45725 45824: gap of unknown length
 * 45825 47844: contig of 2020 bp in length
 * 47845 47944: gap of unknown length
 * 47945 61027: contig of 13083 bp in length
 * 61028 61127: gap of unknown length
 * 61128 63312: contig of 2085 bp in length
 * 63313 63312: gap of unknown length
 * 63313 70560: contig of 7248 bp in length
 * 70561 70660: gap of unknown length
 * 70661 73758: contig of 2998 bp in length
 * 73759 73758: gap of unknown length
 * 73759 103653: contig of 29895 bp in length
 * 103654 103753: gap of unknown length
 * 103754 193735: contig of 89982 bp in length.

FEATURES

source

1..193735
 /organism="Callitrix jacchus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9483"
 /clone="CH259-368E20"
 /clone_1lb="CH259"


```
misc_feature      /note="BAC resource: http://bacpac.chori.org/"
1..29309
/note="assembly_fragment
clone_end:17
vector_side:left"
29410..37712
/note="assembly_fragment"
37813..40472
/note="assembly_fragment"
40573..45724
/note="assembly_fragment"
45825..47844
/note="assembly_fragment"
47945..61027
/note="assembly_fragment"
61128..63212
/note="assembly_fragment"
63313..70560
/note="assembly_fragment"
70661..73658
/note="assembly_fragment"
73759..103653
/note="assembly_fragment"
103754..193735
/note="assembly_fragment
clone_end:SP6
vector_side:right"
127712..193735
/note="clone overlaps with GenBank Accession Number
AC148551 clone CH259-256024 (center project name fvv)"
misc_feature
```

ORIGIN

```
Query Match      59.0%; Score 23.6; DB 2; Length 193735;
Best Local Similarity 76.3%; Pred. No. 60;
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
```

```
Oy      2  GGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db      89109  GGAAATCACTCCAAAATCCCTCAAAATCATGCA 89072
```

```
RESULT 15
AC008897      117026 bp      DNA      linear      PRI 07-NOV-2000
LOCUS      Homo sapiens chromosome 5 clone CTD-2235C13, complete sequence.
AC008897
AC008897.7  GI:11119448
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 117026)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 117026)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (07-NOV-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 7, 2000 this sequence version replaced gi:7711370.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.5.
```

```
STS Content:
WI-9815 G05428
SHGC-16751 G15410
WI-7171 G06431.
FEATURES
Source      Location/Qualifiers
1..117026
```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2235C13"
```

ORIGIN

```
Query Match      57.5%; Score 23; DB 9; Length 117026;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
Oy      1  AGGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db      72634  AGGAACGGATTACAACTATCACCCTTCTCAAGCA 72672
```

```
Search completed: September 16, 2005, 03:01:29
Job time : 139.222 secs
```


This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: September 15, 2005, 08:10:17 ; Search time 19.3656 Seconds
(without alignments)
12227.321 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatactaccctcccaaacatcctcattcattgcaac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_GeneSeq_16Dec04:*
1: geneSeq19808:*
2: geneSeq19909:*
3: geneSeq20008:*
4: geneSeq20018:*
5: geneSeq20018:*
6: geneSeq20028:*
7: geneSeq20028:*
8: geneSeq20038:*
9: geneSeq20038:*
10: geneSeq20038:*
11: geneSeq20038:*
12: geneSeq20048:*
13: geneSeq20048:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40	100.0	40	5	AAH76337	Aah76337 Z. mays M
2	40	100.0	50	5	AAH76336	Aah76336 Z. mays M
3	40	100.0	158	5	AAH76334	Aah76334 Z. mays M
4	40	100.0	235	5	AAH76340	Aah76340 Z. mays M
5	40	100.0	1394	2	AAH07408	Aax07408 Zea mays
6	40	100.0	1394	2	AAH07409	Aax07409 Zea mays
7	40	100.0	1394	5	AAH76332	Aah76332 Z. mays M
8	40	100.0	1394	5	AAH76333	Aah76333 Z. mays M
9	22.4	56.0	226215	11	ACM45146	AcM45146 Human gene
10	21.6	54.0	4110	8	ABZ10202	AbZ10202 Haematopo
11	21.6	54.0	6309	6	ABL12304	AbL12304 Human imm
12	21.6	54.0	7110	10	ADB54282	AdB54282 Pretreat
13	21.6	54.0	7110	10	ADB54196	AdB54196 Human lym
14	21.6	54.0	7110	13	AD889580	Ad889580 Oligonuc
15	21.4	53.5	1341	6	ABZ13750	AbZ13750 Arabidops
16	21.4	53.5	6681	6	ABL21255	AbL21255 Human imm
17	21.4	53.5	6681	6	ABL54304	AbL54304 Chemical
18	21.4	53.5	8168	6	ABL12928	AbL12928 Human imm
19	21.4	53.5	8168	6	AA653328	Aa653328 Chemical
20	21.4	53.5	24401	4	ABL03396	AbL03396 Drosophila

ALIGNMENTS

C	21	21	52.5	2000	12	ADJ41594	AdJ41594 Plant cDN
C	22	21	52.5	2792	12	ADM98941	Adm98941 Diterpene
C	23	21	52.5	3117	12	ADM98942	Adm98942 Diterpene
C	24	21	52.5	5507	8	ABZ10204	AbZ10204 Haematopo
C	25	21	52.5	5507	8	ABZ10058	AbZ10058 Haematopo
C	26	21	52.5	8404	4	AA646500	Aa646500 Tumour su
C	27	21	52.5	8404	6	ABL33595	AbL33595 Human imm
C	28	21	52.5	8404	10	ADB54216	AdB54216 Pretreat
C	29	21	52.5	8404	10	ADB54088	AdB54088 Pretreat
C	30	21	52.5	8404	10	ADB84076	AdB84076 Human lym
C	31	21	52.5	8404	10	ADB84152	AdB84152 Human lym
C	32	21	52.5	8404	13	AD889242	Ad889242 Oligonuc
C	33	21	52.5	8404	13	AD889516	Ad889516 Oligonuc
C	34	21	52.5	10286	4	AA653308	Aa653308 Chemical
C	35	21	52.5	10286	6	ABK28147	AbK28147 DNA trans
C	36	21	52.5	63115	10	ADC85419	AdC85419 Mouse Tle
C	37	21	52.5	63294	9	ADA02939	Ada02939 Mouse Tle
C	38	21	52.5	63294	10	ADB72677	AdB72677 Mouse Tle
C	39	21	52.5	63294	12	ADM74534	AdM74534 Murine ca
C	40	21	52.5	106315	11	ACN43866	AcN43866 Human gen
C	41	20.8	891	8	ABZ51893	AbZ51893 Aspergill	
C	42	20.8	52.0	6620	4	AA645488	Aa645488 Chemical
C	43	20.8	52.0	6620	6	ABK28416	AbK28416 DNA trans
C	44	20.8	52.0	6620	6	ABN80279	AbN80279 Human che
C	45	20.8	52.0	110000	5	AA161373_1	Continuation (2 of

ALIGNMENTS

RESULT 1
AAH76337
ID AAH76337 standard; DNA; 40 BP.
AC AAH76337;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region fragment.
XX
KW Ms45; male tissue; regulatory region; transcription; male fertility;
XX
KW hybrid seed; de.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
XX essential for initiating transcription of the Ms45 gene useful for
XX mediating fertility in a male plant.
XX
PS Claim 14; Page 32; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
XX comprising nucleotide sequences essential for initiating transcription of
XX the Ms45 gene. A method of mediating male fertility in a plant is
XX provided that involves introducing an expression vector comprising a
XX promoter operably linked to (I) into a plant where the exogenous gene
XX impacts male fertility of the plant and (I) controls expression of the
XX exogenous gene. A method of producing hybrid seeds is also provided. The
XX present sequence represents a DNA fragment upstream of the TATA box of a


```
CC 2. mays Ms45 male-tissue preferred regulatory region nucleotide sequence
XX
SQ Sequence 40 BP; 14 A; 14 C; 3 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 40; DB 5; Length 40;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 1 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 40
RESULT 2
AAH76336
ID AAH76336 standard; DNA; 50 BP.
AC AAH76336;
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX 2. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
OS Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -72 to -111 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 50 BP; 18 A; 17 C; 5 G; 10 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 40; DB 5; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 1 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 50
RESULT 3
AAH76334
ID AAH76334 standard; DNA; 158 BP.
```

```
XX
XX AAH76334;
AC
XX 29-OCT-2001 (first entry)
DT
XX
XX 2. mays Ms45 male tissue-preferred regulatory region fragment.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
XX
XX 13-FEB-2001; 2001WO-US004527.
XX
XX 15-FEB-2000; 2000US-00504487.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
XX WPI; 2001-514772/56.
XX
XX A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
XX Claim 5; Page 47; 50pp; English.
XX
XX The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a DNA fragment -38 to -195 bases upstream of
CC the TATA box of a Z. mays Ms45 male-tissue preferred regulatory region
CC nucleotide sequence
XX
XX Sequence 158 BP; 41 A; 50 C; 21 G; 46 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 40; DB 5; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 40
DB 86 AGGATACCTTACTCCCAACATCCATCTTACTCATGCAAC 125
RESULT 4
AAH76340
ID AAH76340 standard; DNA; 255 BP.
AC AAH76340;
XX
XX 29-OCT-2001 (first entry)
DT
XX
XX 2. mays Ms45 promoter fragment.
DE
XX Ms45; male tissue; regulatory region; transcription; male fertility;
KM hybrid seed; promoter; ds.
XX
XX Zea mays.
XX
XX WO200160997-A2.
XX
XX 23-AUG-2001.
```



```
RESULT 7
AAH76332
ID AAH76332 standard; DNA; 1394 BP.
XX
AC AAH76332;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 46; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 311 C; 231 G; 441 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCCCAACATCCATCTTACTGACGCAAC 40
DB 1239 AGGATACCTACTCCCAACATCCATCTTACTGACGCAAC 1278
XX
RESULT 8
AAH76333
ID AAH76333 standard; DNA; 1394 BP.
XX
AC AAH76333;
XX
DT 29-OCT-2001 (first entry)
XX
DE Z. mays Ms45 male tissue-preferred regulatory region encoding DNA.
XX
KM Ms45; male tissue; regulatory region; transcription; male fertility;
KW hybrid seed; ds.
XX
OS Zea mays.
XX
PN WO200160997-A2.
XX
```

```
PD 23-AUG-2001.
XX
PF 13-FEB-2001; 2001WO-US004527.
XX
PR 15-FEB-2000; 2000US-00504487.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Albertsen MC, Fox TW, Garnaat CW, Huffman G, Kendall TL;
XX
DR WPI; 2001-514772/56.
XX
PT A male tissue-preferred regulatory region comprising nucleotide sequences
PT essential for initiating transcription of the Ms45 gene useful for
PT mediating fertility in a male plant.
XX
PS Claim 4; Page 47; 50pp; English.
XX
CC The invention provides a male tissue-preferred regulatory region (I)
CC comprising nucleotide sequences essential for initiating transcription of
CC the Ms45 gene. A method of mediating male fertility in a plant is
CC provided that involves introducing an expression vector comprising a
CC promoter operably linked to (I) into a plant where the exogenous gene
CC impacts male fertility of the plant and (I) controls expression of the
CC exogenous gene. A method of producing hybrid seeds is also provided. The
CC present sequence represents a nucleic acid sequence encoding an Ms45 male
CC -tissue preferred regulatory region from Z. mays
XX
SQ Sequence 1394 BP; 411 A; 309 C; 232 G; 442 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 40; DB 5; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.1e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATACCTACTCCCAACATCCATCTTACTGACGCAAC 40
DB 1239 AGGATACCTACTCCCAACATCCATCTTACTGACGCAAC 1278
XX
RESULT 9
ACN45146
ID ACN45146 standard; DNA; 226215 BP.
XX
AC ACN45146;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human genomic sequence hCG1639824.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1948; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which
```


are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to CarcinoMa Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bioclip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining CarcinoMa Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 226215 BP; 65331 A; 42804 C; 45218 G; 72862 T; 0 U; 0 Other;

Query Match 56.0%; Score 22.4; DB 11; Length 226215;
Best Local Similarity 72.5%; Pred. No. 1.2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1 AGGATACCTACTCCCAACATCCATCTTACTCATGCAAC 40
183206 AGAATTCCTCTCTCCAAACACAGCATCTTACTCTTTAC 183245

RESULT 10
AB210202/C
ID AB210202 standard; DNA; 4110 BP.

AB210202;

16-JUN-2003 (first entry)

Haematopoietic cell proliferation disorder related DNA sequence #342.

Human; haematopoietic cell proliferation disorder; cytostatic;
Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
cytosine methylation state; gene; db.

Homo sapiens.

MO20027722-A2.

03-OCT-2002.

26-MAR-2002; 2002MO-EP003401.

26-MAR-2001; 2001US-0278333P.

(EPig-) EPIGENOMICS AG.

Berlin K, Braun A, Dierker J, Guetig D, Howe A, Mueller J;

Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu B;

Lewin A, Lipscher B, Mater S, Model F, Mueller V, Otto T, Pellet C;

Schweppe I, Ziebarth H;

WPI, 2003-018942/01.

Detecting and differentiating between hematopoietic cell proliferative

disorders, comprises contacting a target nucleic acid with a reagent that

distinguishes between methylated and non-methylated CpG dinucleotides.

Claim 28; SEQ ID NO 342; 117bp; English.

The present invention describes a method for detecting and differentiating between haematopoietic cell proliferative disorders associated with at least 1 gene and/or their regulatory regions in a subject. The method comprises contacting a target nucleic acid in a biological sample obtained from the subject with at least 1 reagent, which distinguishes between methylated and non-methylated CpG dinucleotides within the target nucleic acid. AB209861 to AB211118 represent specifically claimed nucleotide sequences from the present invention. Oligonucleotides from the present invention can be used: for

differentiating between healthy haematopoietic cells and proliferative disorder haematopoietic cells; for differentiating between acute lymphocytic leukaemia and acute myelogenous leukaemia; as probes for determining the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) of haematopoietic cell proliferation disorder related sequences and their complements; and as primers for the amplification of haematopoietic cell proliferation disorder related DNA sequences. The nucleotide sequences from the present invention can also be used for detecting a predisposition to, differentiation between CC haematopoietic cell proliferation disorders, treatment and/or monitoring of CC haematopoietic cell proliferation disorders. The present method enables a highly specific classification of haematopoietic cell proliferative disorders allowing for improved and informed treatment of patients

Sequence 4110 BP; 945 A; 0 C; 1132 G; 2033 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 8; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

4 ATACCTACTCTCCCAACATCCATCTTAC 31
3083 AACCTAACCCCAACATCCATCTTAC 3056

RESULT 11
ABL32304/C
ID ABL32304 standard; DNA; 6309 BP.

ABL32304;

26-MAR-2002 (first entry)

Human immune system associated gene SEQ ID NO: 277.

Human; immune system associated; cytosine methylation; antiasthmatic;
antiarteriosclerotic; antihaemic; cyrostatic; noctropic;
neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
antihaematic; antiarthritic; antidiabetic; antipsoriatic;
antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;

db.

Homo sapiens.

MO200200928-A2.

03-JAN-2002.

02-JUL-2001; 2001WO-EP007537.

30-JUN-2000; 2000DE-01032529.

01-SEP-2000; 2000DE-01043826.

(EPig-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K;

WPI, 2002-130909/17.

Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation.

Claim 1; SEQ ID NO 277; 32bp + Sequence Listing; German.

The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 6309 BP; 1758 A; 69 C; 1364 G; 3118 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 6; Length 6309;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAAGATCCATCTTACATGCAA 39
DB 3335 ATACCTATTCCTCATGACCAATCTCATATCCAA 3300

RESULT 12
ADB54282/C
ID ADB54282 standard; DNA; 7110 BP.

XX ADB54282;
XX
DT 04-DEC-2003 (first entry)
XX

DE Pretreated genomic DNA region 206.

XX colon cell proliferative disorder; non methylated CpG dinucleotide;
XX cytostatic; cancer; adenoma; carcinoma; cytosine methylation state; ds.

XX Unidentified.

XX WO2003072821-A2.

XX 04-SEP-2003.

XX 27-FEB-2003; 2003MO-EP002035.

XX 27-FEB-2002; 2002EP-00004551.

XX (EPIG-) EPIGENOMICS AG.

XX Adorjan P, Burger M, Maier S, Nimrich I, Becker E, Lesche R;
PI Rujan T, Schmitt A;

XX MPI; 2003-731620/69.

PT Detecting and differentiating between colon cell proliferative disorders
PT associated with a gene or its regulatory regions comprises contacting a
PT target nucleic acid in a biological sample obtained from the subject with
PT a reagent.

XX Claim 32; SEQ ID NO 338; 74pp; English.

CC The invention relates to a novel method for detecting and differentiating
CC between colon cell proliferative disorders associated with at least one
CC gene or its regulatory regions. The method comprises contacting a target
CC nucleic acid in a biological sample obtained from the subject with at
CC least one reagent or a series of reagents, where the reagent or series of
CC reagents, distinguishes between methylated and non methylated CpG
CC dinucleotides within the target nucleic acid. The molecules of the
CC invention demonstrate cytosine activity whilst the method may useful
CC for detecting and differentiating between colon cell proliferative
CC disorders, including cancers such as colon adenoma and colon carcinoma.
CC The PNA (peptide nucleic acid)-oligomers are useful as probes for
CC determining cytosine methylation state or single nucleotide
CC polymorphisms. The current sequence is that of the pretreated genomic DNA
CC region of the invention. This sequence is not shown within the
CC specification but is taken from Wipoweb.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAAGATCCATCTTAC 31
DB 3083 AAACCTAACCCAAACATCCATCTTAC 3056

RESULT 13
ADB84196/C
ID ADB84196 standard; DNA; 7110 BP.

XX ADB84196;
XX

DT 29-JAN-2004 (first entry)
XX

DE Human lymphoid cell proliferative disorder gene derived DNA #132.

XX ds; lymphoid cell proliferative disorder; methylation;
XX methylated CpG dinucleotide; single nucleotide polymorphism; SNP;
XX diffuse large B-cell lymphoma; mantle cell lymphoma;
XX chronic lymphocytic leukemia; small lymphocytic lymphoma;
XX follicular lymphoma; diagnosis; prognosis.

XX Homo sapiens.

XX WO2003044226-A2.

XX 30-MAY-2003.

XX 25-NOV-2002; 2002MO-EP013265.

XX 23-NOV-2001; 2001DE-01057491.

XX 28-DEC-2001; 2001DE-01064501.

XX (EPIG-) EPIGENOMICS AG.

XX Burger M, Caldwell C, Genc B, Becker E, Maier S, Nimrich I;
PI MPI; 2003-457621/43.

PT Detecting and differentiating between lymphoid cell proliferative
PT disorders comprises contacting a target nucleic acid with at least one
PT reagent that distinguishes between methylated and non-methylated CpG
PT dinucleotides.

XX Claim 26; SEQ ID NO 192; 448pp; English.

CC The invention relates to a method of detecting and differentiating
CC between lymphoid cell proliferative disorders associated with at least
CC one gene and/or their regulatory regions in a subject by contacting a
CC target nucleic acid in a biological sample obtained from the subject with
CC at least one reagent or series of reagents that distinguish between
CC methylated and non-methylated CpG dinucleotides within the target nucleic
CC acid. The genes and/or their regulatory regions are preferably selected
CC from MRL1, CSNK2B, EGR4, AR, CDK4, RB2, CDC25A, GRB beta, MYO1, CDH3,
CC MYC1, ELK1, ABL1, APC, BCL2, CDH1, CDKN1A, CDKN1B, CDKN2A, CDKN2B, FOS,
CC GSTR1, HIC-1, MGMT, MLH1, MOS, MYC, PTEN, RBL2, TGFBR2, TP73, CDKN1C,
CC GSK3beta, ESRR1, APAF1, BAK1, BAX or HMOX5. Oligomers, peptide nucleic
CC acid (PNA)-oligomers and/or isolated nucleic acids based on the sequences
CC of the genes are useful for detecting the methylation state of all the
CC CpG dinucleotides within one or more the sequences, or their complements,
CC for determining the cytosine methylation state and or single nucleotide
CC polymorphisms (SNPs), and for differentiating at least two of the medical
CC conditions such as diffuse large B-cell lymphoma, mantle cell lymphoma,
CC chronic lymphocytic leukemia, small lymphocytic lymphoma and follicular
CC lymphoma. They are also useful for detecting of a predisposition to,
CC differentiation between subclasses, diagnosis, prognosis, treating and/or
CC monitoring of lymphoid cell proliferative disorder. This sequence
CC represents a nucleic acid of a pretreated genomic DNA derived from the
CC above mentioned genes.

XX Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 10; Length 7110;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;

Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACTTCTCTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 14

AD89580/C
 ID AD89580 standard; DNA; 7110 BP.

AC ADS89580;

DT 18-NOV-2004 (first entry)

DE Oligonucleotide of the invention SEQ ID NO:596.

KM 89; cell proliferative disorder; breast; methylation; cytostatic;
 KM gene therapy; single nucleotide polymorphism; SNP.

OS Unidentified.

PN WO2004035803-A2.

PD 29-APR-2004.

PF 01-OCT-2003; 2003WO-EP010881.

PR 01-OCT-2002; 2002DE-01045779.

PR 07-JAN-2003; 2003DE-01000096.

PR 17-APR-2003; 2003DE-01017955.

PA (EPIC-) EPIDENOMICS AG.

PI Foekens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
 PI Nimmrich I, Rujan T, Schmitt A, Schmitt M, Look MP, Marx A;

DR WPI; 2004-348466/32.

PT Predicting responsiveness of a subject with breast cell proliferative
 PT disorder; useful for treating or differentiating breast cell
 PT proliferative disorders comprises analyzing methylation pattern of a
 PT genomic DNA from the subject.

PS Disclosure; SEQ ID NO 596; 104pp; English.

CC The invention relates to a novel method for predicting the responsiveness
 CC of a subject with a cell proliferative disorder of the breast tissues to
 CC a therapy comprising analyzing the methylation pattern of a target
 CC nucleic acid by contacting at least one of the target nucleic acids in a
 CC biological sample obtained from the subject prior to or during treatment.
 CC The method of the invention has cytostatic activity, and may have a use
 CC in gene therapy. The set of oligonucleotides comprising at least two of
 CC the oligomers are useful for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
 CC methods, nucleic acid, oligonucleotide, and kit are useful for the
 CC treatment, characterization, classification and/or differentiation, of
 CC breast cell proliferative disorders. The method is also useful for
 CC predicting the responsiveness of a subject with a cell proliferative
 CC disorder of the breast tissues to a therapy. The present sequence is used
 CC in the exemplification of the invention.

SQ Sequence 7110 BP; 1818 A; 0 C; 1721 G; 3571 T; 0 U; 0 Other;

Query Match 54.0%; Score 21.6; DB 13; Length 7110;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACTTCTCTCCCAACATCATCTTAC 31
 DB 3083 AACCTAACCCCAACATCATCTTAC 3056

RESULT 15

AB213750/C
 ID AB213750 standard; DNA; 1341 BP.

AC AB213750;

DT 21-JAN-2003 (first entry)

DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1555.

KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

PN WO200216655-A2.

PD 28-FEB-2002.

PF 24-AUG-2001; 2001WO-US026685.

PR 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-0300111P.

PA (SCRI) SCRIPPS RES INST.

PI (SYGN) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

DR WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and
 PT producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 1555; 577pp + Sequence Listing; English.

CC The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (AB213750-AB217574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

SQ Sequence 1341 BP; 460 A; 231 C; 319 G; 331 T; 0 U; 0 Other;

Query Match 53.5%; Score 21.4; DB 6; Length 1341;
 Best Local Similarity 71.8%; Pred. No. 1.5e+02;
 Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCTCCCAACATCATCTTACATCCAA 39
 DB 288 AGTCACCTATTCTCCCAACATCATCTTACATCCAA 250

Search completed: September 15, 2005, 21:33:17
 Job time : 22.3656 secs

This Page Blank (up to)

This Page Blank (up to)

Query Match	100.0%	Score 40;	DB 3;	Length 1394;
Best Local Similarity	100.0%	Pred. No. 4.6e-07;		
Matches 40;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

TYPE: DNA
ORGANISM: Human
US-09-949-016-65770

Query Match
Best Local Similarity 71.1%; Score 20.4; DB 4; Length 601;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATCCTACTCCCAACATCCATCTTACTCATGCA 39
Db 414 GGATCCTGCGGCTCAGCAATCCTCTGCGCTCAGGCA 377

RESULT 6
US-09-774-528-184
Sequence 184, Application US/09774528
Patent No. 6743619
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyen
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonghong
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Wang, Dunrui
APPLICANT: Dermanac, Radoje T.
TITLE OF INVENTION: No. 6743619el Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 802
CURRENT APPLICATION NUMBER: US/09/774,528
CURRENT FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 441
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 184
LENGTH: 973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (231)..(749)
US-09-774-528-184

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 4; Length 973;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATCCTACTCCCAACATCCATCTTACTCATGCA 39
Db 626 GGAATCACTCCAAAGAAACCTCGAATCATGCA 663

RESULT 7
US-08-945-056-4
Sequence 4, Application US/08945056
Patent No. 6077994
GENERAL INFORMATION:
APPLICANT: Coupland, George M.
APPLICANT: Puterill, Joanna J.
TITLE OF INVENTION: Genetic control of flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderye PC
STREET: 8th Floor, 1100 No. 6077994th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (RPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,056
FILING DATE: 20-OCT-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02561
FILING DATE: 01-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9422083.7
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4201 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
STRAIN: landberg erecta
POSITION IN GENOME:
MAP POSITION: chromosome 5
US-08-945-056-4

Query Match
Best Local Similarity 51.0%; Score 20.4; DB 3; Length 4201;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GATACCTACTCCCAACATCCATCTTACT 32
Db 2631 GATACCAAGCTCCACACATCAATCTTACT 2660

RESULT 8
US-09-949-016-13668/C
Sequence 13668, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13668
LENGTH: 51770
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(51770)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13668

Query Match
51.0%; Score 20.4; DB 4; Length 51770;

Best Local Similarity 71.1%; Pred. No. 1.5e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db 9851 GGATTCCTGGCTCAGCATCCTCCTCGCTCAGGCAA 9814

RESULT 9
US-09-949-016-17420

; Sequence 17420, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17420

; LENGTH: 84571

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(84571)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17420

Query Match

Best Local Similarity 51.0%; Score 20.4; DB 4; Length 84571;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db 24559 GGAAATCAACTCCAAAAGAACCTTCAAAACATGCAA 24596

RESULT 10
US-09-949-016-11824/c

; Sequence 11824, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 11824

; LENGTH: 126200

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(126200)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-11824

Query Match 51.0%; Score 20.4; DB 4; Length 126200;
Best Local Similarity 71.1%; Pred. No. 1.9e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db 53133 GGAAATCAACTCCAAAAGAACCTTCAAAATCATGCAA 53096

RESULT 11

US-09-949-016-13193/c

; Sequence 13193, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13193

; LENGTH: 126200

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(126200)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-13193

Query Match

Best Local Similarity 51.0%; Score 20.4; DB 4; Length 126200;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCCATCTTACTCATGCAA 39
Db 53133 GGAAATCAACTCCAAAAGAACCTTCAAAATCATGCAA 53096

RESULT 12
US-09-610-040-5

; Sequence 5, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Moesener, Jeffrey

; APPLICANT: Gorslach, Jorn

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloci, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS

; FILE REFERENCE: 9128.14

; CURRENT APPLICATION NUMBER: US/09/610,040

; PRIOR FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 5099

; TYPE: DNA


```

; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (4045)..(4049)
; OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).
US-09-610-040-5

Query Match
Best Local Similarity 88.0%; Score 20.2; DB 3; Length 5099;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY
Db 2855 ACCTACCCCAACAATCCATCTTA 30
ACCTACCCCAACAATCCATCTTA 30

RESULT 13
US-10-267-763-5
; Sequence 5, Application US/10267763
; Patent No. 6800459
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Azenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DI1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5099
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(5099)
; OTHER INFORMATION: n = A,T,C or G
US-10-267-763-5

Query Match
Best Local Similarity 88.0%; Score 20.2; DB 4; Length 5099;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY
Db 2855 ACCTACCCCAACAATCCATCTTA 30
ACCTACCCCAACAATCCATCTTA 30

RESULT 14
US-09-949-016-88655
; Sequence 88655, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0010307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88655
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-88655

Query Match
Best Local Similarity 50.0%; Score 20; DB 4; Length 601;
Matches 26; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY
Db 293 GTATAGTCTCCCAACAACGTTGCTTATATATAAA 330
GGATACCTACTCCCAACAATCCATCTTACTATGCA 39

RESULT 15
US-09-620-312D-22
; Sequence 22, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aildong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 22
; LENGTH: 2615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(849)
US-09-620-312D-22

Query Match
Best Local Similarity 72.2%; Score 20; DB 4; Length 2615;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY
Db 1855 ATACGATTCACCAACAACATCTCTCATGGAA 1890
ATACGATTCACCAACAACATCTCTCATGGAA 39

Search completed: September 15, 2005, 08:25:10
Job time : 9.05048 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 15, 2005, 07:10:11 ; Search time 26.0505 Seconds
(without alignments)
10230.248 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40

Sequence: 1 aggatcaccctccccaacacatccatccatcgcgaac 40

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 333185599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/prodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/prodata/2/pubpna/US10J_NEW_PUB.seq:*
- 23: /cgn2_6/prodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/prodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	40	20	US-10-713-381-6
2	40	100.0	50	20	US-10-713-381-5
3	40	100.0	158	20	US-10-713-381-3
4	40	100.0	255	20	US-10-713-381-9
5	40	100.0	1394	20	US-10-713-381-1
6	40	100.0	1394	20	US-10-713-381-2
7	24	60.0	85779	20	US-10-713-993-6880

8	23.2	58.0	706	18	US-10-424-599-31006	Sequence 31006, A
9	22.6	56.5	294	20	US-10-425-115-143168	Sequence 143168, A
10	22.4	315	18	US-10-424-599-83179	Sequence 83179, A	
11	22.4	56.0	226215	13	US-10-087-192-1948	Sequence 1948, Ap
12	21.6	54.0	4110	20	US-10-473-126-342	Sequence 342, App
13	21.6	54.0	6309	15	US-10-311-455-277	Sequence 277, App
14	21.4	53.5	1341	9	US-09-938-8428-1555	Sequence 1555, Ap
15	21.4	53.5	1341	11	US-09-938-8428-1555	Sequence 1555, Ap
16	21.4	53.5	6681	15	US-10-311-455-128	Sequence 128, App
17	21.4	53.5	6681	16	US-10-240-452-4	Sequence 4, Appl1
18	21.4	53.5	8168	15	US-10-311-455-901	Sequence 901, Appl
19	21.4	53.5	8168	18	US-10-240-454-23	Sequence 23, Appl
20	21.4	53.5	3673778	16	US-10-312-841-1	Sequence 1, Appl1
21	21.4	53.5	3673778	16	US-10-312-841-2	Sequence 2, Appl1
22	21	293	18	US-10-424-599-24043	Sequence 24043, A	
23	21	52.5	931	18	US-10-424-599-20320	Sequence 20320, A
24	21	52.5	2000	17	US-10-260-238-2594	Sequence 2594, Ap
25	21	52.5	2792	18	US-10-041-018-361	Sequence 361, App
26	21	52.5	3117	18	US-10-041-018-362	Sequence 362, App
27	21	52.5	5507	20	US-10-473-126-198	Sequence 198, App
28	21	52.5	5507	20	US-10-473-126-344	Sequence 344, App
29	21	52.5	8404	15	US-10-311-455-1568	Sequence 1568, Ap
30	21	52.5	8404	18	US-10-221-714A-222	Sequence 222, App
31	21	52.5	10286	14	US-10-239-676-13	Sequence 13, Appl
32	21	52.5	10286	15	US-10-240-453-21	Sequence 21, Appl
33	21	52.5	63294	11	US-09-997-722-205	Sequence 205, App
34	21	52.5	106315	13	US-10-087-192-178	Sequence 178, App
35	21	52.5	2140405	13	US-10-027-632-76212	Sequence 76212, A
36	21	52.5	2140405	17	US-10-027-632-76212	Sequence 76212, A
37	20.8	52.0	2263	19	US-10-437-963-4616	Sequence 4616, Ap
38	20.8	52.0	3292	17	US-10-027-632-112652	Sequence 112652, Ap
39	20.8	52.0	3292	17	US-10-027-632-112652	Sequence 112652, Ap
40	20.8	52.0	6620	14	US-10-239-676-186	Sequence 186, App
41	20.8	52.0	6620	15	US-10-240-453-290	Sequence 290, App
42	20.8	52.0	513509	10	US-09-754-853A-4	Sequence 4, Appl1
43	20.6	51.5	970	18	US-10-424-599-91073	Sequence 91073, A
44	20.6	51.5	1103	22	US-10-499-352A-170	Sequence 170, App
45	20.6	51.5	1889	22	US-10-499-352A-178	Sequence 178, App

ALIGNMENTS

RESULT 1
US-10-713-381-6
; Sequence 6, Application US/10713381
; Publication No. US2004022131A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-6

Query Match 100.0%; Score 40; DB 20; Length 40;
Best Local Similarity 100.0%; Pred.No. 3.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGATCCTACCTCCCAACATCCATCTTACATCCAC 40

Db 1 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 40

RESULT 2

US-10-713-381-5
; Sequence 5, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-5

Query Match 100.0%; Score 40; DB 20; Length 50;
Best Local Similarity 100.0%; Pred. No. 3.3e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 40
Db 11 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 50

RESULT 3

US-10-713-381-3
; Sequence 3, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 158
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-3

Query Match 100.0%; Score 40; DB 20; Length 158;
Best Local Similarity 100.0%; Pred. No. 4.2e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 40
Db 86 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 125

RESULT 4

US-10-713-381-9
; Sequence 9, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 255
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-9

Query Match 100.0%; Score 40; DB 20; Length 255;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 40
Db 99 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 138

US-10-713-381-1
; Sequence 1, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.
; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUFFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; TITLE OF INVENTION: USING SAME
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-1

Query Match 100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 40
Db 1239 AGGATACCTACTCCCAAAACATCCATCTTACTCATGCAC 1278

US-10-713-381-2
; Sequence 2, Application US/10713381
; Publication No. US20040221331A1
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, MARC C.


```

; APPLICANT: FOX, TIMOTHY W.
; APPLICANT: GARNAT, CARL W.
; APPLICANT: HUPFMAN, GARY
; APPLICANT: KENDALL, TIMMY L.
; TITLE OF INVENTION: MALE TISSUE-PREFERRED REGULATORY REGION AND METHOD OF
; FILE REFERENCE: 578R
; CURRENT APPLICATION NUMBER: US/10/713,381
; CURRENT FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: 08/880,499
; PRIOR FILING DATE: 1997-06-23
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Zea mays
US-10-713-381-2

Query Match          100.0%; Score 40; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 6,4e-06;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACAATCATCTTACTCATGCAAC 40
DB 1239 AGGATACCTACTCCCAACAATCATCTTACTCATGCAAC 1278

RESULT 7
US-10-719-993-6880/c
; Sequence 6880, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6880
; LENGTH: 85779
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (1)...(85779)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6880

Query Match          60.0%; Score 24; DB 20; Length 85779;
Best Local Similarity 84.4%; Pred. No. 38;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 ATACTACTCCCAACAATCATCTTACTCAT 35
DB 16082 AAACCTACTCCCAACAATCATCTTACTCAT 16051

RESULT 8
US-10-424-599-31006
; Sequence 31006, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 31006
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-31006
```

```

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 31006
; LENGTH: 706
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128002C.1
US-10-424-599-31006

Query Match          58.0%; Score 23.2; DB 18; Length 706;
Best Local Similarity 77.8%; Pred. No. 31;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACAATCATCTTACTCATGCAAC 40
DB 455 TCCCTTCTCCCAAAACCTCATCAACCTCATCAAC 490

RESULT 9
US-10-425-115-143168
; Sequence 143168, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 143168
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62049C.1
US-10-425-115-143168

Query Match          56.5%; Score 22.6; DB 20; Length 294;
Best Local Similarity 86.2%; Pred. No. 45;
Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACTACTCCCAACAATCATCTTACTCAT 32
DB 226 ATACTGTCCTCATCATCTTACTCAT 254

RESULT 10
US-10-424-599-83179
; Sequence 83179, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 83179
; LENGTH: 315
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46127C.1
US-10-424-599-83179
```


Query Match 56.0%; Score 22.4; DB 18; Length 315;
Best Local Similarity 81.2%; Pred. No. 55;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 35
DB 211 AACGAACCTCTCAACATCATCTTACTCAT 242

RESULT 11
US-10-087-192-1948
; Sequence 1948, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:

; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1948
; LENGTH: 226215
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1948

Query Match 56.0%; Score 22.4; DB 13; Length 226215;
Best Local Similarity 72.5%; Pred. No. 2e+02;
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
DB 183206 AGAATTCCTGCTTCCCAACACAGCATCTTACTCTTTTAC 183245

RESULT 12
US-10-473-126-342/c

; Sequence 342, Application US/10473126
; Publication No. US20040234973A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
; SEQ ID NO 342
; LENGTH: 4110
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-342

Query Match 54.0%; Score 21.6; DB 20; Length 4110;
Best Local Similarity 85.7%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCAT 31
DB 3083 AACCTTACCCCAACATCATCTTACTCAT 3056

RESULT 13

US-10-311-455-277/c
; Sequence 277, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBERG, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 277
; LENGTH: 6309
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-277

Query Match 54.0%; Score 21.6; DB 15; Length 6309;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCAAC 39
DB 3335 ATACCTACTCCCAACATCATCTTACTCATGCAAC 3300

RESULT 14
US-09-938-842A-1555/c
; Sequence 1555, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1555
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1555

Query Match 53.5%; Score 21.4; DB 9; Length 1341;
Best Local Similarity 71.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCAAC 39
DB 288 AGTCCATCTTCCCAACATCTCTGTGCTTACACAA 250

RESULT 15

Page 5

Page 5

Page 5

Page 5

Page 5

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 15, 2005, 16:07:32 ; Search time 124.754 Seconds

(without alignments)
12204.542 Million cell updates/sec

Title: US-10-713-381-1_COPY_1239_1278

Perfect score: 40
Sequence: 1 aggatacctaccctcccaaacatccatctcactcgcacac 40

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Lifting first 45 summaries

Database :

EST: *
1: gb_ests1.*
2: gb_ests2.*
3: gb_hnc.*
4: gb_ests3.*
5: gb_ests4.*
6: gb_ests5.*
7: gb_ests6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	100.0	687	9	CC656939 OGMDO20TV
2	40	100.0	915	9	CG224225 OGIAG08TV
3	40	100.0	963	9	CC656933 OGMDO20TV
4	23.8	59.5	528	7	H24137 ym50g08.r1
5	23.6	59.0	375	8	B2706605 SM41681-G
6	23.2	58.0	738	7	CO117589 GR_EB01H
7	23.2	58.0	783	9	CR181951 Revere
8	23.2	58.0	1018	9	CL109241 ISB1-SID1
9	23.2	58.0	1253	6	CD495749 CDA18-D05
10	23	57.5	405	8	AZ883408 RPCI-23-1
11	23	57.5	798	4	BG506830 601861443
12	23	57.5	968	4	BP143726 601789606
13	22.8	57.0	771	4	BM406466 EST580781
14	22.6	56.5	410	2	AM251178 EST305315
15	22.6	56.5	560	2	BE205271 EST397947
16	22.6	56.5	624	8	CR843833 GROMAA78D
17	22.6	56.5	624	8	B2173201 CH230-369
18	22.6	56.5	644	9	CR843905 GROMAA79A
19	22.6	56.5	664	8	AZ400686 IM0167P01
20	22.6	56.5	691	2	AM687128 NF006C08R
21	22.6	56.5	736	9	CR818130 GROMAA43D
22	22.6	56.0	1080	9	CL058859 CH216-88P
23	22.4	56.0	298	2	BB264116 BB264116
24	22.4	56.0	940	7	CK864514 AGENCOURT

C	25	22.4	56.0	940	7	CK871857	CK871857	AGENCOURT
C	26	22.4	56.0	951	7	CN317130	CN317130	AGENCOURT
	27	22.2	55.5	273	1	A1206216	A1206216	q227E06.X
	28	22.2	55.5	346	1	A1698201	A1698201	wag7911.X
	29	22.2	55.5	398	1	AA725064	AA725064	a106h07.s
	30	22.2	55.5	417	1	A1467879	A1467879	cj78e09.X
	31	22.2	55.5	479	2	BF601210	BF601210	266128 MA
	32	22.2	55.5	491	1	A1382034	A1382034	te33g05.X
C	33	22.2	55.5	609	4	BM027586	BM027586	GTT000109
	34	22.2	55.5	727	7	CK949745	CK949745	4074916 B
C	35	22.2	55.5	802	9	CG811817	CG811817	FSAM45TR
C	36	22.2	55.5	940	9	CG383910	CG383910	OGZAP61TV
	37	22.2	55.5	1090	9	CL084441	CL084441	ISB1-3G16
C	38	22	55.0	207	7	CN313825	CK913825	p3fmgcf.0
C	39	22	55.0	330	3	CNS09857	CK917651	Single re
	40	22	55.0	467	3	CG982704	CG982704	CH240_164
C	41	22	55.0	478	3	CNS0986U	BX047050	Single re
	42	22	55.0	506	6	CNS09070	BX036328	Single re
C	43	22	55.0	506	8	AZ021739	AZ021739	RPCI-23-3
C	44	22	55.0	548	3	CNS09071	BX036329	Single re
	45	22	55.0	576	8	AQ321478	AQ321478	RPCI11-10

ALIGNMENTS

RESULT 1	CC656939	687 bp	DNA	linear	GSS 19-JUN-2003
LOCUS	OGMDQ20TV.ZM 0.7 1.5 KB	Zea mays	genomic clone	ZMM8MA0554D15,	
DEFINITION	genomic survey sequence.				
ACCESSION	CC656939	GI:32060231			
VERSION	CC656939.1				
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE	White, J.A., Fraser, C.M., Buddiman, M.A., Bedell, J.A., Rohlfing, T., Resnick, A., Nurnberg, A., Robbins, D. and Lakey, N.				
AUTHORS	White, J.A., Fraser, C.M., Buddiman, M.A., Bedell, J.A., Rohlfing, T., Resnick, A., Nurnberg, A., Robbins, D. and Lakey, N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: OGMDO20TV				
	Contact: Cathy Whitelaw				
	TIGR				
	9712 Medical Center Drive, Rockville, MD 20850, USA				
	Tel: 301-838-5843				
	Fax: 301-838-0208				
	Email: whitelaw@tigr.org				
	Seq primer: TF				
	Class: sheared ends.				
FEATURES	Location/Qualifiers				
source	1..687				
	/organism="Zea mays"				
	/mol_type="genomic DNA"				
	/strain="B73"				
	/db_xref="taxon:4577"				
	/clone="ZMM8MA0554D15"				
	/clone_1ib="ZM_0.7 1.5 KB"				
	/note="Vector: pBCK-j Site_1: HincII, 0.7-1.5 kb				
	methylation filtered genomic DNA library"				
ORIGIN					
Query Match	100.0%; Score 40; DB 9; Length 687;				
Best Local Similarity	100.0%; Pred. No. 3.2e-05;				
Matches	40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
1	AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 40				
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
Db	337 AGGATACCTACTCCCAACATCCATCTTACTCATGCAC 298				

RESULT 2
CG224225 915 bp DNA linear GSS 22-AUG-2003
DEFINITION OGIAG08TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0716B15,
genomic survey sequence.
ACCESSION CG224225
VERSION CG224225.1 GI:34124113
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 915)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
TITLE JOURNAL
COMMENT Other_GSSs: OGIAG08TV
Contact: Cathy WhiteJaw
TIGR

FEATURES
source
1..915
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0716B15"
/note="Vector: pBCSK-, Site 1; HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 915;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 40
|||||
Db 618 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 657
|||||

RESULT 3
CC656933 963 bp DNA linear GSS 19-JUN-2003
LOCUS CC656933
DEFINITION OGMDQ20TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMA0554D15,
genomic survey sequence.
ACCESSION CC656933
VERSION CC656933.1 GI:32060225
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 963)
WhiteJaw,C.A., Quackenbush,J., Van Aken,S., Uteerback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Numborg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
TITLE JOURNAL
COMMENT Other_GSSs: OGMDQ20TV
Contact: Cathy WhiteJaw
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whiteJaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..963
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBMA0554D15"
/note="Vector: pBCSK-, Site 1; HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 100.0%; Score 40; DB 9; Length 963;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 40
|||||
Db 836 AGGATACCTACTCCCAACATCCATCTTACTATGCAAC 875
|||||

RESULT 4
H24137 528 bp mRNA linear EST 06-JUL-1995
LOCUS YMS0908.r1 Soares infant brain IN1B Homo sapiens cDNA clone
DEFINITION IMAGE:51939 5', mRNA sequence.
ACCESSION H24137
VERSION H24137.1 GI:892832
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 528)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maixra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
TITLE JOURNAL
COMMENT Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1161
Source: IMAGE Consortium, LLNL
High quality sequence strops: 349
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1161 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 349.
Location/Qualifiers
1..528
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:424875"
/db_xref="taxon:9606"
/clone_lib="IMAGE:51939"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares infant brain IN1B"
/note="Organ: whole brain; Vector: latmid BA; Site_1: Not

I, Site 2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' ACTGGAAGATTGCGCGCGAGAAATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the lacZmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Patricia Bonaldo."

Query Match 59.5%; Score 23.8; DB 7; Length 528;
Best Local Similarity 80.0%; Pred. No. 97;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 ATACTTCTCCCAACATCTTACTATGCA 38
346 AATCTACTCTCAACATCTTACTATGCA 380

RESULT 5
BZ706605/c 375 bp DNA linear GSS 18-MAY-2003
LOCUS SM416B1-G07.55.13.ab1 Spider Monkey genomic BAC library Ateles
DEFINITION Geoffroyi genomic, genomic survey sequence.
ACCESSION BZ706605
VERSION BZ706605.1 GI:30843142
KEYWORDS GSS.
SOURCE Ateles geoffroyi (black-handed spider monkey)
ORGANISM Ateles geoffroyi
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae;
Ateles.

REFERENCE 1 (bases 1 to 375)
AUTHORS Qian, Y., Jin, L. and Su, B.
TITLE BAC end sequences of spider monkey genomic library
JOURNAL Unpublished (2003)
COMMENT Contact: Qian Y., Jin L., Su B.
Center for Genome Information
University of Cincinnati
Kettering Lab, 3223 Eden Ave., Cincinnati, OH 45267-0056, USA
Tel: 1-513-558-6678
Fax: 1-513-558-4505
Email: subhucmail.uc.edu
Seq primer: T7
Class: BAC ends.

FEATURES

source

1..375
/organism="Ateles geoffroyi"
/mol_type="genomic DNA"
/db_xref="black-handed spider monkey"
/db_xref="taxon:9509"
/sex="Male"
/tissue_type="Skin"
/cell_type="Fibroblast"
/cell_line="AG05352"
/dev_stage="3 DA"
/clone_lib="Spider Monkey genomic BAC library"
/note="Vector: PBACe3.6; Site 1: EcoRI; Genomic DNA was partially digested with EcoRI. Vector, PBACe3.6, recombinants were transformed into DH10B."

ORIGIN

Query Match 59.0%; Score 23.6; DB 8; Length 375;
Best Local Similarity 74.4%; Pred. No. 1.1e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATCTACTCTCCCAACATCTTACTATGCA 40
255 GGATCTACTCTGGAACATCTTACTATGCA 217

RESULT 6
COL17589
COL17589

LOCUS COL17589 738 bp mRNA linear EST 16-JUN-2004
DEFINITION GR_EB01H15.r GR_Eb Gossypium raimondii cDNA clone GR_EB01H15 3', mRNA sequence.
ACCESSION COL17589
VERSION COL17589.1 GI:48816276
KEYWORDS EST.

SOURCE Gossypium raimondii
ORGANISM Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eustosids II; Malvales; Malvaceae; Malvoideae; Gossypium.

REFERENCE 1 (bases 1 to 738)
AUTHORS Kim, H., Yu, Y., Kudrna, D., Hatfield, J., Stum, D., Mueller, C., and Udall, J. A., Rapp, R. A., Wendel, J. F., Rao, K., Soderlund, C., and Wang, R. A.

TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu

FEATURES
source
Location/Qualifiers
1..738
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_EB01H15"
/tissue_type="floral"
/dev_stage="3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCW.SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by invitrogen with RNA supplied by Wendie Lab. Directional cloned into NotI-EV. Clones plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 58.0%; Score 23.2; DB 7; Length 738;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACTTCTCCCAACATCTTACTATGCA 39
224 ATACACATCTCCATACACGCAAGTACTATGCA 259

RESULT 7
CRI81951 783 bp DNA linear GSS 06-JUL-2004
LOCUS Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN51p21, genomic survey sequence.
DEFINITION CRI81951
VERSION CRI81951.1 GI:49960800
KEYWORDS GSS; genome survey sequence; MICEP.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 783)
AUTHORS Adams, D. J., Biggs, P. J., Cox, A. V., Davies, R. M., van der Weyden, L., Jonkers, J., Smith, J., Plumb, R. W., Taylor, R. G., Nishijima, I., Yu, Y., Rogers, J., and Bradley, A.

TITLE Direct Substitution
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICEP
LOCATION/Qualifiers
1..783
/organism="Mus musculus"
/mol_type="genomic DNA"

/db_xref="taxon:10090"
/clone="MHPN361p21"
/clone_lib="MHPN"

ORIGIN

Query Match 58.0%; Score 23.2; DB 9; Length 783;
Best Local Similarity 77.8%; Pred. No. 1.8e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 4 ATACCTACTCCCAACATCATCTTACTCATGCA 39
|||||
514 ATACCTACTCTAAGCCATGATCATCTTACCGAGAA 549

RESULT 8

CL109241/c 1018 bp DNA linear GSS 05-JAN-2004
LOCUS ISB1-51D14.Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-51D14,
DEFINITION genomic survey sequence.
ACCESSION CL109241
VERSION CL109241.1 GI:40602876
KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 1018)
Kremetzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E., and Wilson, R.

A physical map of the xenopus tropicalis genome
Unpublished (2003)
Contact: Richard K Wilson

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu

Insert Length: 75000 Std Error: 0.00
Seq primer: Sp6 ATTAGTGACACTATAG
Class: BAC ends

High quality sequence start: 72
High quality sequence stop: 524.
Location/Qualifiers

FEATURES

source

1..1018
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"

/db_xref="taxon:8364"
/clone="ISB1-51D14"

/clone_lib="ISB1"

/note="Vector: pBelobAC11, ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Query Match 58.0%; Score 23.2; DB 9; Length 1018;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACATCATCTTCTCATGCAAC 40
|||||
Db 156 TACATCTCTCCATACATCATCTTTATGAAATPAC 121

RESULT 9

CD495749 1253 bp mRNA linear EST 12-JUN-2003
LOCUS CD495749
DEFINITION CD495749
CD495749
CD495749
CD495749.1 GI:31422780

EST.
Gasterosteus aculeatus (three spined stickleback)
Gasterosteus aculeatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
Gasterosteidae; Gasterosteus.
1 (bases 1 to 1253)
Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M.,
Schmitz, J., and Myers, R.M.

Expressed sequence tags from Gasterosteus aculeatus
Unpublished (2003)
Contact: Kingsley, DM

HMT and Department of Developmental Biology
Stanford University School of Medicine
Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA
Tel: 650 725 5954
Fax: 650 725 7739

Email: kingsley@cgm.stanford.edu
Plate: 18
High quality sequence start: 15
High quality sequence stop: 667.
Location/Qualifiers

FEATURES

source

1..1253

/organism="Gasterosteus aculeatus"

/mol_type="RNA"

/strain="Salinas river, CA"

/db_xref="taxon:69293"

/clone="CD495749"

/sex="mixed male and female"

/tissue_type="heads and internal organs combined"

/dev_stage="adult"

/clone_lib="SHGC-CDA"

/note="Vector: Lambda ZAP Express/PBK-CMV, Site_1: EcoRI
(5' adaptor); Site_2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-cDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dt sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of PBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual PBK-CMV phagemid clones for
EST sequencing."

ORIGIN

Query Match 58.0%; Score 23.2; DB 6; Length 1253;
Best Local Similarity 77.8%; Pred. No. 2e+02;
Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACATCATCTTACTCATGCAAC 40
|||||
Db 1047 TACCTCTTCCCAACATCATCTTACCATGCAAC 1082

RESULT 10

AZ883408 405 bp DNA linear GSS 05-MAR-2001
LOCUS RPCI-23-18912.TJ RPCI-23 Mus musculus genomic clone RPCI-23-18912,
DEFINITION genomic survey sequence.
ACCESSION AZ883408
VERSION AZ883408.1 GI:13202353
KEYWORDS GSS.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 405)
Zhao, S., Nierman, W., Feldblum, T., Malek, J., Shatman, S.,
Akinret, B., Levins, M., McGinn, S., Tesgaye, G., Geer, K., Krol, M., de
Jong, P., and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other GSS: RPCI-23-18912.TV
Contact: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ezhah@tigr.org
Clones are derived from the mouse BAC library RPCT-23. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tc/bac_ends/mouse/bac_end_intro.html
Plate: 189 row: L column: 2
Seg primer: SP6
Class: BAC ends.

FEATURES

SOURCE

Location/Qualifiers
1..405
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCT-23-189L2"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 57.5%; Score 23; DB 8; Length 405;
Best Local Similarity 74.4%; Pred. No. 1.9e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGGATACCTACTCCCAACATCATCTTACTCATGCA 39
DB 182 AGGTACCAAAATCCCAACATGCTTCTGTATCATGCA 144

RESULT 11
BGS06830/c 798 bp mRNA linear EST 27-MAR-2001
LOCUS 601861443F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4071154 5',
DEFINITION mRNA sequence.
ACCESSION BGS06830
VERSION BGS06830.1 GI:13468347
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 798)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM916 row: e column: 11
High quality sequence stop: 2.
Location/Qualifiers

FEATURES

SOURCE

1..798
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="IMAGE:4071154"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccatcgcc); Site_2: SfiI (ggccatcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCATTTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCATG-dT(30)BN-3' (where B = A, C, G or N and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

ORIGIN

Query Match 57.5%; Score 23; DB 4; Length 798;
Best Local Similarity 83.9%; Pred. No. 2.2e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 10 ACTCCCAACATCATCTTACTCATGCAAC 40
DB 222 ACTCCCAACAGCCTTCTTATGATGATC 192

RESULT 12
BF143726 968 bp mRNA linear EST 24-OCT-2000
LOCUS 601789606F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4020600 5',
DEFINITION mRNA sequence.
ACCESSION BF143726
VERSION BF143726.1 GI:10982766
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 968)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLM9274 row: k column: 01
High quality sequence start: 3
High quality sequence stop: 648.
Location/Qualifiers

FEATURES

SOURCE

1..968
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:4020600"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu30"
/note="Organ: Lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; transgenic model MMT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 57.5%; Score 23; DB 2; Length 968;
Best Local Similarity 74.4%; Pred. No. 2.3e+02;
Matches 29; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GGATACCTACTCCCAACATCATCTTACTCATGCAAC 40
 BM406466
 LOCUS 818 GGCTCCGCTACTCCCAACACACCCACCCACACCCAC 856

RESULT 13
 BM406466 771 bp mRNA linear EST 10-MAR-2003
 LOCUS EST580781 potato roots Solanum tuberosum cDNA clone cPRO27F14 5'
 DEFINITION end, mRNA sequence.

ACCESSION BM406466
 VERSION BM406466.1 GI:18258084
 KEYWORDS EST.
 SOURCE Solanum tuberosum (potato)
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanales; Solanum.
 1 (bases 1 to 771)
 van der Hoeven, R., Sun, H., Karamycheva, S.A., Tsai, J., Van Aken, S., Utechtback, T., Chiemingo, A., Bougri, O., Buel, C.R., Romning, C., Tanksley, S. and Baker, B.
 Generation of ESTs from potato roots
 Unpublished (2001)
 Contact: Robin Buel
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL:
 http://genome.arizona.edu/orders/
 Seq primer: T3.

FEATURES
 source
 Location/Qualifiers
 1..771
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cPRO27F14"
 /tissue_type="roots"
 /dev_stage="in vitro grown stem cuttings"
 /lab_host="SOLR"
 /clone_lib="potato roots"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Cornell University, Tanksley lab.
 sequencing: The Institute for Genomic Research. Roots were isolated from in vitro grown stem cuttings on CM medium. Roots were isolated two weeks after placing the stem cuttings from in vitro grown plants on medium."

ORIGIN
 Query Match 57.0%; Score 22.8; DB 4; Length 771;
 Best Local Similarity 79.4%; Pred. No. 2.6e+02;
 Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TACCTACTCCCAACATCATCTTACTCATGCA 38
 BM406466
 LOCUS 649 TCCCTCCTTCACACATCACCATCTCAAGCA 682

RESULT 14
 AM257178 410 bp mRNA linear EST 20-DEC-1999
 LOCUS EST305315 KV2 Medicago truncatula cDNA clone KV2-7D15, mRNA
 DEFINITION sequence.
 ACCESSION AM257178
 VERSION AM257178.1 GI:6605435
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE 1 (bases 1 to 410)
 Vandenbosch, K., Hurt, J., Moore, J., Beremand, P., Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Frazer, C.M.
 ESTs from roots of Medicago truncatula after Rhizobium inoculation
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenbosch.umn.edu
 Texas A&M EST name: T115883e
 TIGR sequence name: MTAAO207K
 More information is available at: . (and for clone ordering info)
 http://chryslr.tamu.edu/medicago
 Seq primer: SKmod (CPA gAA CTA gtc gAT CC).

FEATURES
 source
 Location/Qualifiers
 1..410
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="KV2-7D15"
 /tissue_type="Seedling roots"
 /dev_stage="2 days post-inoculation with Sinorhizobium meliloti"
 /lab_host="E. coli strain SOLR"
 /clone_lib="KV2"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN
 Query Match 56.5%; Score 22.6; DB 2; Length 410;
 Best Local Similarity 86.2%; Pred. No. 2.8e+02;
 Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 6 ACCTACTCCCAACATCATCTTACTCA 34
 BM205271
 LOCUS 16 ACCAAGCTCCCAACATCATCTGCTCA 44

RESULT 15
 BE205271 560 bp mRNA linear EST 05-SEP-2000
 LOCUS EST397947 KV0 Medicago truncatula cDNA clone pKV0-21123, mRNA
 DEFINITION sequence.
 ACCESSION BE205271
 VERSION BE205271.1 GI:8748567
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
 1 (bases 1 to 560)
 Vandenbosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Frazer, C.M.
 ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 Contact: Vandenbosch K
 Department of Plant Biology
 University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738

Email: kvandend@cds.umn.edu

Texas A&M University name: T265043e

TIGR sequence name: MTGBK60TK

More information is available at:

<http://chryslr.tamu.edu/medicago>

Seq primer: Skmod (CTA gaa CTA gtc gat CC).

FEATURES

source

```
1..560
  location/Qualifiers
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /cultivar="genotype A17"
    /db_xref="taxon:3880"
    /clone="pKV0-21123"
    /tissue_type="Seedling roots"
    /dev_stage="Immediately prior to inoculation with
Sinorhizobium meliloti (0 hour)"
    /lab_host="E.coli strain XL0LR"
    /clone_id="KV0"
    /note="Vector: Bluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant Lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."
```

ORIGIN

Query Match

Best Local Similarity 86.2%; Score 22.6; DB 2; Length 560;

Matches 25; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

6 ACCTACTCCCAACATCATCTTACTCA 34

DB

24 ACCAATCCCAACATCATCTGCTCA 52

Search completed: September 16, 2005, 08:08:40
 Job time : 131.754 secs

This Page Blank (uspto)